LOGIQUE

Cours de deuxième année de bachelier en Philosophie

P. Gribomont

2003-2008

Table des matières

Avant-propos

La logique tient une part importante dans les études de philosophie. Une raison naturelle en est que la logique est une branche de la philosophie ; d'Aristote à Wittgenstein en passant par Leibniz et Kant, elle a toujours été considérée comme telle. Bien plus, en dehors de son intérêt intrinsèque, la logique peut être utilisée, comme outil et aussi comme objet d'étude, dans d'autres domaines de la philosophie, comme l'épistémologie et la philosophie du langage. Sur ce plan, les logiques formelles sont particulièrement intéressantes ; elles constituent des langages dont le pouvoir d'expression est grand mais dont la syntaxe, la sémantique et la pragmatique restent simples, du moins si l'on compare avec les langages naturels. L'étude formalisée de la logique est une discipline mathématique,[1] donc scientifique, dont les développements suscitent beaucoup de questions intéressant le philosophe. Naturellement, les autres sciences, telles la géométrie, la physique et la biologie, suscitent également de telles questions mais, d'un point de vue didactique et pédagogique, l'etude de la logique formalisée est attrayante car elle peut être abordée avec profit sans connaissances spécifiques préalables.

Pour réfléchir à propos de la logique, il est indispensable d'en maîtriser d'abord les bases. Dès l'enfance, l'être humain développe ses aptitudes logiques, dans le cadre de l'école et en dehors. Cet apprentissage est le plus souvent inconscient et ne concerne que la logique non formelle, mais il n'en est pas moins efficace. Le cours de première candidature a systématisé ces acquis et les a enrichis.

Le cours de logique de deuxième candidature développe la logique classique et, plus spécifiquement, la logique des propositions et celle des prédicats. La logique comporte diverses techniques, étayées par des développements théoriques. Nous introduisons relativement peu de théorie et de techniques ; en fait, nous nous limitons à l'essentiel. En revanche, le cours propose un important éventail d'exercices, dont la résolution implique une parfaite assimilation des concepts et des techniques.

Les exercices de logique sont essentiellement de trois types. En premier lieu, l'étudiant(e), ayant acquis un bagage de connaissance suffisant, peut utiliser la logique, formelle ou non, pour clarifier un problème impliquant du raisonnement et le résoudre. La difficulté des exercices de ce genre réside souvent plus dans le passage du langage naturel vers la logique que dans la démarche logique elle-même. En deuxième lieu, la logique formelle, comme toute discipline scientifique, donne naissance à une grande variété d'exercices impliquant la mise en œuvre de ses méthodes et de ses techniques. Les exercices de ce type sont parfois routiniers ; ils résistent rarement à une approche méthodique. Leur principal intérêt est de raffermir, de concrétiser et parfois de préciser dans l'esprit de l'apprenant les divers concepts et méthodes qui lui sont proposés. Enfin, il existe des exercices plus difficiles, plus abstraits, qui consistent à réfléchir sur les concepts et méthodes de la logique elle-même, et à se poser à leur propos les questions "comment ?" et — surtout — "pourquoi ?". Ces exercices peuvent graduellement conduire à aborder avec profit des questions plus générales, sortant du cadre de la logique ; nous en rencontrerons quelques exemples.

[1] Ce point est largement incontestable ; le fait que des philosophes tels que Leibniz, Frege, Whitehead, Russell, Wittgenstein, von Wright et beaucoup d'autres aient contribué de manière substantielle au développement de la logique formelle montre simplement que le philosophe est souvent amené à faire œuvre de mathématicien . . . et qu'il n'y a pas d'opposition entre l'esprit de finesse et l'esprit de géométrie.

Les exercices de logique visent plus à approfondir la connaissance déjà acquise qu'à l'augmenter, mais quelques exceptions seront bienvenues ! Concrètement, nous restreindrons notre champ de réflexion à la logique des propositions et à celle des prédicats du premier ordre ; l'expérience montre d'ailleurs clairement qu'il vaut mieux maîtriser convenablement ces logiques classiques avant d'aborder, notamment, les logiques modales.

Les notes qui suivent se veulent un résumé succinct mais autonome de la logique formelle élémentaire. Cet exposé est, nous l'espérons, sauvé de l'aridité par les exercices, les discussions et les réflexions qui l'émaillent, et qui sont peut-être la partie la plus importante. Nous avons essayé d'éviter les développements trop compliqués et trop chargés de formalisme ; nous avons surtout voulu proscrire tout passage peu clair, puisque "It is a safe rule to apply that, when a mathematical or philosophical author writes with a misty profundity, he is talking nonsense" [A.N. Whitehead, An Introduction to Mathematics, 1911 ; Oxford Univ. Press. 1990 paperback, ISBN 0-19-500211-3]. En revanche, ces notes ne dissimulent pas le caractère mathématique de la logique, puisque "To create a healthy philosophy you should [. . .] be a good mathematician" [B. Russell, 1935 ; cité dans E.T, Bell, Men of Mathematics, Simon and Schuster, New York, 1937].

1 Introduction

1.1 Enseigner la logique formelle

Du point de vue de son enseignement, la logique formelle élémentaire se trouve dans une situation paradoxale. D'une part, cet enseignement est favorisé par plusieurs facteurs objectifs mais, d'autre part, les résultats obtenus sont souvent décevants. Nous développons ici brièvement ces deux points, et proposons quelques pistes pour améliorer la situation.

1.1.1 Quelques atouts

Citons d'abord trois raisons pour lesquelles un cours d'introduction à la logique formelle devrait être un cours facile à donner, et facile à assimiler.

– *La matière proprement dite est objectivement facile.*
Analyser une formule propositionnelle, telle $((p \land q) \Rightarrow r) \equiv (p \Rightarrow (q \Rightarrow r))$, est nettement plus simple qu'analyser une formule arithmétique ou algébrique telle que $\sqrt{ab} \leq (a+b)/2$. En effet, les propositions ne peuvent être que vraies ou fausses, tandis que les nombres forment un ensemble infini. Cela a pour conséquence que les opérations, les règles et les méthodes de la logique propositionnelle sont moins nombreuses et plus simples que celles de l'algèbre élémentaire. D'une manière analogue, il est plus facile d'analyser une formule du calcul des prédicats, telle que (exemple classique) $\forall x\,(P(x) \Rightarrow Q(x)) \Rightarrow (\forall x\,P(x) \Rightarrow \forall x\,Q(x))$, que de résoudre une équation intégrale, telle que $y(x) = 1 + \int_0^x y(t)\,\mathrm{d}t$. Enfin, comme nous le verrons, presque tous les résultats de la logique élémentaire sont des exercices simples et s'établissent de manière quasi systématique, alors que chaque théorème de mathématique élémentaire, fût-il aussi vieux que celui de Pythagore, ressemble à un défi.
– *Les références de bonne qualité, accessibles à l'autodidacte, ne manquent pas.*
Même si, à l'échelle de la philosophie et de la mathématique, la logique formelle est une branche plutôt jeune, elle a quand même plus d'un siècle ; le plus récent résultat que nous verrons, le principe de résolution, date de 1965 (il est même nettement antérieur en ce qui concerne la logique propositionnelle). En mathématique, tous les domaines de base ont fait l'objet de présentations didactiques nombreuses et soignées ; la logique n'échappe pas à la règle. De plus, la logique formelle ne s'est pas développée à partir de rien ; le calcul des propositions doit beaucoup aux Stoïciens et le calcul des prédicats est issu de la théorie du syllogisme d'Aristote et des logiciens du Moyen-Age ; cette théorie reste parfaitement lisible aujourd'hui.
– *Les mathématiques préparent à la logique.*
La logique est la science du raisonnement et de l'expression formelle du raisonnement. Tout étudiant est amené à raisonner et à exprimer le fruit de ses cogitations oralement ou par écrit . . . Bien plus, les écueils traditionnels de la logique élémentaire (implication, démonstration, variables libres et liées, etc.) ont déjà été rencontrés dans l'enseignement secondaire, dans des contextes mathématiques souvent plus difficiles. En particulier, la notion d'implication formalise le lien existant entre l'hypothèse d'un théorème et sa thèse ; distinguer les rôles des variables x et y dans la formule $\forall x\,P(x, y)$ n'est pas plus

difficile que distinguer les rôles de t et x dans l'intégrale $\int_0^x f(t)\,\mathrm{d}t$, ou ceux de i et j dans $\sum_i A_{ji}x_i$.

1.1.2 Quelques problèmes ...

Pourquoi alors l'étudiant, reconnaissant rapidement et sans hésitation la validité des formules $((p \wedge q) \Rightarrow r) \equiv (p \Rightarrow (q \Rightarrow r))$ et $\sqrt{ab} \leq (a+b)/2$, hésitera-t-il devant des questions innocentes, telles que

> *Soient A, B, X et Y des formules quelconques.*
> *On pose $A' =_{def} (X \Rightarrow (A \Rightarrow Y))$ et $B' =_{def} (X \Rightarrow (B \Rightarrow Y))$.*
> *Si $A \Rightarrow B$ est vrai, que peut-on dire de $B \Rightarrow A$, de $A' \Rightarrow B'$ et $B' \Rightarrow A'$?*

et

> *Quel lien logique y a-t-il entre les formules*
> *$\forall x\, \forall y\, (P(x) \Rightarrow Q(y))$ et $\exists x\, P(x) \Rightarrow \forall x\, Q(x)$?*

Nous n'avons naturellement pas d'explication définitive à ce problème, et encore moins de remède infaillible, mais on peut néanmoins noter que les trois points cités plus haut, bien qu'objectivement favorables, ne sont pas dépourvus d'effets pervers.

La facilité de la matière peut susciter trois types de réactions négatives. Tout d'abord, "si c'est trop simple, ce n'est pas utile". Les applications non triviales de la logique sont pourtant nombreuses, mais le temps manque parfois pour les aborder. Ensuite, et cela surtout à propos de la logique propositionnelle, "pourquoi vouloir formaliser et théoriser à propos d'une arithmétique simpliste, limitée à 0 (faux) et 1 (vrai) ?". Enfin, la facilité conduit à l'imprudence, qui elle-même mène à l'erreur ! La logique renferme quand même quelques pièges ...

Les bons livres existent, sans aucun doute, mais ne correspondent pas toujours aux attentes et besoins du lecteur. Un simple exposé du type "hypothético-déductif" habituellement utilisé en mathématique ne convient pas, même si paradoxalement la logique élémentaire s'y prête très bien. Ce genre d'exposé se lit avec peu d'effort mais conduit seulement à une compréhension passive et superficielle des concepts. En outre, un tel exposé ne donne pas de justification à l'existence même de la logique mathématique et ne fera qu'amplifier les réactions négatives évoquées plus haut.[2]

Notons enfin que son bagage mathématique, s'il aide objectivement l'étudiant à aborder le présent cours, pourrait aussi contribuer à susciter sa méfiance ... Est-il naïf d'espérer que ce cours contribue à réconcilier le lecteur avec la démarche mathématique ?

1.1.3 Quelques solutions

Les remèdes existent. Une approche historique et philosophique des concepts [Gochet et Gribomont, 1989] est un excellent moyen de contrer les réactions négatives, en montrant que beaucoup d'efforts ont été nécessaires pour aboutir aux concepts simples et épurés sur lesquels se base la logique moderne. Elle montre aussi que les progrès réalisés au cours des siècles l'ont souvent été à l'occasion de problèmes concrets ; on voit enfin que la formalisation de l'expression des raisonnements a été la voie royale conduisant à une meilleure compréhension

[2]Un exposé de nature mathématique est cependant très utile à l'apprenant, dès qu'il a maîtrisé les bases et acquis une certaine pratique.

de ceux-ci. L'inconvénient de cette approche est qu'elle allonge grandement la taille de l'exposé, surtout si on le complète d'une introduction à des problèmes extra-logiques [Gochet et Gribomont, 1994, 2000] auxquels la logique apporte une solution partielle ou complète. Tout en restant persuadé de l'intérêt pédagogique d'une telle approche de la logique, nous devons admettre qu'elle est peu compatible avec la durée de 30 heures prévue au programme, surtout pour des auditeurs fortement sollicités par ailleurs.

Un moyen radical de balayer les objections de simplicité et d'inutilité est de dépasser la matière reconnue comme indispensable et d'aborder quelques grands problèmes tels l'incomplétude et l'indécidabilité de la plupart des théories, et les limites de la mécanisabilité du raisonnement, un sujet qui a toujours passionné les logiciens, depuis Leibniz jusqu'à Quine, en passant par Gödel et Herbrand. On obtient alors un cours d'allure nettement mathématique, plutôt volumineux et difficile, dont l'introduction dans un curriculum de candidature en philosophie serait malaisée à justifier.

Il semble donc que les problèmes liés à l'enseignement de la logique se résolvent surtout par des développements supplémentaires, dont le simple volume peut rebuter l'étudiant. En dépit de cette inquiétante inflation, on constate que la partie centrale de la logique peut s'exposer en relativement peu de temps, pour un profit intellectuel considérable à condition de respecter une stricte discipline. Le point crucial de cette discipline est que l'étudiant doit assimiler la logique élémentaire comme l'arithmétique élémentaire ; il doit pouvoir "doubler" le raisonnement méthodique et rigoureux par une compréhension intuitive des formules. Il doit arriver, par exemple, à *rejeter* l'énoncé de logique (incorrect !)

> *Si $A \Rightarrow B$ est vrai, alors $(X \Rightarrow (A \Rightarrow Y)) \Rightarrow (X \Rightarrow (B \Rightarrow Y))$ est vrai.*

aussi rapidement que l'énoncé algébrique (incorrect !)

> *Si $a \leq b$ est vrai, alors $(y-a)-x \leq (y-b)-x$ est vrai.*

En mathématique, il est extrêmement pénible de mémoriser des démonstrations vues comme des textes linéaires dont tous les mots ont la même importance. Il est de loin préférable d'associer à un théorème un objet concret (au sens large) à partir duquel on peut reconstituer aisément la démonstration du théorème.

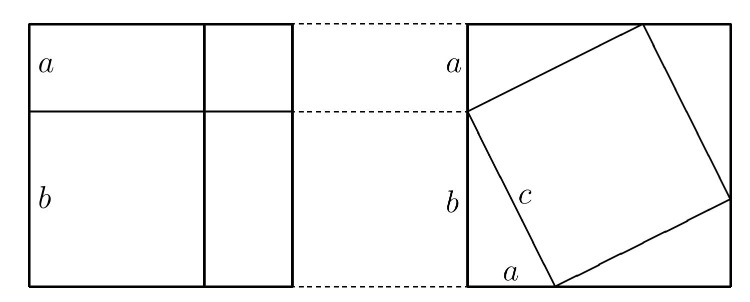


FIG. 1 – Clef du théorème de Pythagore.

Le dessin de gauche de la figure 1 comporte deux carrés intérieurs dont les dimensions sont a et b ainsi que deux rectangles de côtés a et b. Ce dessin illustre notamment la formule

$(a+b)^2 = a^2 + 2ab + b^2$. Le dessin de droite comporte un carré intérieur de dimension c, ainsi que quatre triangles rectangles de petits côtés a et b et d'hypoténuse c. L'aire totale des deux rectangles étant égale à celle des quatre triangles, l'aire totale $a^2 + b^2$ des deux carrés intérieurs à gauche est égale à l'aire c^2 du carré intérieur à droite. Cette dernière égalité est le théorème de Pythagore.

Ce genre d'objet (ici, une paire de dessins) est naturellement très utile ; il rend évident le théorème auquel il se rapporte. L'inconvénient est qu'il n'est pas facile de découvrir l'objet qui éclairera un résultat important. C'est cependant moins difficile en logique formelle qu'en algèbre ou en arithmétique ; montrer comment ces objets peuvent être découverts et utilisés sera l'un de nos objectifs. Nous essayerons aussi d'importer en logique l'expérience et l'intuition acquises en arithmétique élémentaire.

1.1.4 Digression : Pythagore

Il est difficile, à propos de Pythagore, de distinguer entre faits historiques avérés et légendes invérifiables, mais une tradition constante lui attribue (ou à son école) plusieurs contributions majeures, depuis le concept même de philosophie jusqu'à la découverte de l'influence positive des voyages sur la formation de la jeunesse, en passant évidemment par ses célèbres théorèmes. Nous avons déjà évoqué celui du triangle rectangle ; un autre affirme que $\sqrt{2}$ est irrationnel, c'est-à-dire que $\sqrt{2}$ n'est égal à aucun quotient de deux entiers.

Ces résultats sont très importants mais beaucoup plus significatif est le fait qu'ils aient été démontrés. Avant Pythagore (et, en bien des lieux, longtemps encore après lui), les mathématiques se réduisaient à un ensemble de résultats plus ou moins précis et structurés, tenant de la recette de cuisine[3] ou de la vérité d'inspiration divine,[4] mais jamais démontrés. C'est le concept de preuve qui fonde les mathématiques, la logique mathématique, la physique mathématique et toutes les sciences exactes. Les sciences sont exactes parce qu'elles ne sont pas absolues ; les vérités mathématiques sont subordonnées aux axiomes ayant permis de les démontrer.

Un autre pilier des sciences exactes est la généralité, concept également prôné par Pythagore. Les preuves de l'Ecole pythagoricienne font implicitement usage de la règle de généralisation, sur laquelle nous auront l'occasion de revenir. Par exemple, l'égalité de Pythagore $a^2 + b^2 = c^2$ est établie pour un triangle rectangle sans propriété particulière, donc elle est valable pour tous les triangles rectangles. Le progrès de la science est inséparable du progrès de la généralité. Un exemple célèbre est celui des lois de Kepler, qui décrivent la trajectoire elliptique des planètes autour du soleil. Ces lois ont cependant été éclipsées par la loi de la gravitation universelle imaginée par Newton, car, logiquement, les lois de Kepler ne sont que des corollaires de la loi de Newton. Dans le même ordre d'idée, la théorie du syllogisme catégorique d'Aristote et des Scolastiques reste actuelle par sa rigueur (non formelle) mais a été éclipsée par la logique moderne des prédicats, nettement plus générale.

Revenons enfin au concept de preuve, central en mathématique depuis Pythagore. Trop souvent, l'idée même de preuve formelle fait peur aux débutants, et tout au plus ceux-ci peuvent-ils se résigner à en ingurgiter un certain nombre ; quant à imaginer pouvoir construire eux-mêmes une preuve ... Il est vrai que certaines preuves sont extrêmement difficiles, et se font attendre pendant trois siècles, comme celle du dernier théorème de Fermat que nous évoquerons brièvement plus loin.[5] Cependant, dans beaucoup de cas, la preuve est un sous-produit d'une bonne compréhension des questions étudiées. L'étape cruciale est souvent de se poser la bonne question. On attribue à l'Ecole pythagoricienne la découverte du dodécaèdre régulier, solide dont les douze faces sont des pentagones réguliers égaux. Prouver que ce dodécaèdre existe est facile, penser à s'interroger sur son existence, ou sur celles des polyèdres réguliers convexes en général, l'est moins.[6] Un exemple plus simple : est-ce que $\sqrt{2}$ peut s'écrire sous la forme d'une fraction p/q, où p et q sont des entiers positifs premiers entre eux ? La réponse, négative, semble avoir été une déception pour Pythagore, mais la démonstration (par l'absurde) n'est pas difficile. Supposons que la fraction existe. On a donc $p^2 = 2q^2$ donc p^2 est pair et p aussi.[7] On a donc $p = 2p'$ et $4p'^2 = 2q^2$, d'où q^2 et q sont pairs, d'où p et q ne sont pas premiers entre eux.

Dans ce cours, les preuves difficiles sont rares ... mais savoir ce que l'on démontre est un impératif permanent !

1.2 Qu'est-ce que la logique ?

La logique est la science du raisonnement et de l'expression précise du raisonnement. Tout être humain raisonne et est donc concerné par la logique. Raisonner, c'est produire de l'information à partir d'information préexistante et de certains mécanismes de transformation de l'information. Le mot "produire" est ici ambigu, comme l'illustre une scène célèbre du "Rhinocéros" de Ionesco, où il apparaît que, dans un certain sens, la logique ne "produit" rien et ne fait que restituer ce qui lui est fourni. En fait, le raisonnement est proche du calcul, qui lui aussi transforme l'information. Etant donné un triangle dont la base et la hauteur sont de 6 cm (information préexistante), on sait que l'aire du triangle est de 18 cm^2 ("nouvelle" information). Le mécanisme de production est la règle classique $S = (B \times H)/2$. L'information produite est-elle réellement nouvelle ? Peut-être pas, mais en éprouver de la déception serait reprocher à la logique et aux mathématiques ce qui fait leur valeur, à savoir l'universalité : tout triangle dont la base et la hauteur sont de 6 cm a une aire de 18 cm^2 ; l'information que le calcul produit n'est donc pas originale ou inattendue, mais elle n'en est pas moins utile et pertinente.

L'analogie entre la logique et l'arithmétique est féconde et nous aurons l'occasion d'y revenir. Leibniz évoquait déjà l'intérêt d'un langage et d'une méthode permettant un raisonnement mécanisé, de nature calculatoire ; cela aurait permis, notamment, de remplacer

[3] Mille ans avant Pythagore, les Egyptiens et les Babyloniens connaissaient des techniques permettant de résoudre des problèmes intéressants, mais il s'agissait de recettes étayées uniquement par l'expérience.

[4] La tradition attribue à Pythagore le concept de nombre parfait. Un nombre est parfait s'il est égal à la somme de ses diviseurs propres. Deux exemples sont 6 = 1+2+3 et 28 = 1+2+4+7+14. Saint-Augustin aurait suggéré que le monde a été créé en six jours parce que 6 est un nombre parfait.

[5] Rappelons à cette occasion que Pierre de Fermat (1601-1665) fit une brillante carrière ... de juriste. Les mathématiques, et notamment la théorie des nombres, étaient pour lui un hobby.

[6] Y a-t-il des polyèdres réguliers dont les faces ont n côtés ? On sait que l'angle du n-gone régulier est de $a =_{def} 180(n-2)/n$ degrés, et que si k faces ($k \geq 3$, nécessairement) se rejoignent à un sommet, il faut que $k * a$ soit strictement inférieur à 360 degrés. Pour le triangle équilatéral, $a = 60$, d'où $k \in \{3,4,5\}$; pour le carré, $a = 90$, d'où $k = 3$, pour le pentagone régulier, $a = 108$, d'où $k = 3$; pour l'hexagone régulier, $a = 120$, d'où il n'y a pas de solution pour $n \geq 6$. Il y a donc maximum cinq polyèdres réguliers convexes, et il suffit d'essayer de les construire pour constater qu'il y en a exactement cinq.

[7] Le carré d'un nombre pair est pair, le carré d'un nombre impair est impair.

des controverses stériles et agitées par de froids calculs devant les résultats desquels l'unanimité se ferait immanquablement.

1.2.1 La logique des propositions

La logique la plus simple est celle des propositions. Il s'agit bien d'un calcul, dans lequel les objets ne sont pas les nombres et les expressions numériques, mais les valeurs de vérité ("vrai" et "faux") et les propositions et formules susceptibles d'être vraies ou fausses. Voici un exemple typique de ce calcul. L'information préexistante comporte deux énoncés :

Pour sortir sous la pluie, je prends mon parapluie.
Je suis dehors sans parapluie.

Le mécanisme de calcul, ou plutôt de déduction, est le suivant

$$\frac{A \Rightarrow B\,,\ \neg B}{\neg A}$$

"Si A implique B est vrai, et si B est faux, alors A est faux." L'information nouvelle que l'on peut obtenir ici est "Il ne pleut pas". On a *instancié* A en "il pleut" et B en "je sors muni d'un parapluie".
On notera l'emploi de la convention habituelle : la ligne horizontale sépare les *prémisses* d'un raisonnement (au-dessus de la ligne) et sa *conclusion* (au-dessous de la ligne).

Logique et arithmétique. Dans le raisonnement précédent, le calcul proprement dit est extrêmement simple. C'est une arithmétique des plus rudimentaires, où on ne dispose que de deux "nombres", notés **F** (faux) et **V** (vrai), ou encore, pour parfaire l'analogie, 0 et 1. Les tables correspondant au conditionnel et aux quelques autres opérations logiques seront donc bien plus simples que la table de multiplication par exemple, puisque les tables logiques ne comportent que deux entrées. Une difficulté de la logique par rapport à l'arithmétique est le caractère informel du langage utilisé (le français). L'information représentée par B peut être écrite "je sors muni d'un parapluie" ; on a donc implicitement admis que la phrase "Pour sortir sous la pluie, je prends mon parapluie." est synonyme de "Il pleut implique je sors muni d'un parapluie". On conçoit aisément que, dans des raisonnements plus élaborés, une hypothèse de ce type peut rapidement devenir douteuse.[8] Toutefois, ce problème est du ressort de la linguistique et nous ne l'aborderons pas ici. Plus précisément, nous ne considérerons pas de raisonnement dont la formalisation ne soit élémentaire, voire même déjà faite. L'objet de la logique mathématique sera donc la représentation et l'analyse des raisonnements formalisés. A titre d'exemple, considérons les tables de la négation et du conditionnel :

x	$\neg x$
V	**F**
F	**V**

x	y	$x \Rightarrow y$
V	**V**	**V**
V	**F**	**F**
F	**V**	**V**
F	**F**	**V**

[8] Même dans notre exemple, certains problèmes surgissent. En particulier, il peut avoir commencé à pleuvoir après que je sois sorti (sans parapluie) ; je peux aussi égarer mon parapluie en cours de route.

Ces tables permettent d'établir la validité du mécanisme de raisonnement utilisé au paragraphe précédent. Par simple combinaison, on obtient immédiatement la table ci-dessous :

A	B	$A \Rightarrow B$	$\neg B$	$\neg A$
V	**V**	**V**	**F**	**F**
V	**F**	**F**	**V**	**F**
F	**V**	**V**	**F**	**V**
F	**F**	**V**	**V**	**V**

Valider le mécanisme de raisonnement utilisé dans notre exemple (c'est le "Modus Tollens") consiste à vérifier que, dans tous les cas où les deux prémisses $A \Rightarrow B$ et $\neg B$ sont vraies, la conclusion $\neg A$ est également vraie. Dans les trois premières lignes du tableau, l'une des prémisses est fausse. Ces lignes correspondent à des cas où le Modus Tollens ne s'applique pas. La quatrième ligne correspond au cas où les deux prémisses sont vraies, c'est-à-dire au cas où le mécanisme de raisonnement étudié s'applique ; on observe que la conclusion est également vraie, ce qui achève la vérification.

1.2.2 La logique prédicative

Certains types de raisonements ne peuvent s'analyser au niveau de la proposition ; il faut alors décomposer celle-ci, par exemple sous la forme sujet-prédicat. Un exemple classique est le raisonnement suivant :

- Tous les hommes sont mortels.
- Socrate est un homme.
- Donc, Socrate est mortel.

Dans ce raisonnement, de la famille des syllogismes catégoriques, le "donc" est indubitablement pertinent, mais le codage purement propositionnel ne le montre pas. En effet, les trois propositions contenues dans le syllogisme sont élémentaires et distinctes, ce qui fait de ce raisonnement une instance du schéma

$$\frac{p\,,\ q}{r}$$

Ce schéma est clairement non valide. En revanche, en logique prédicative, ce même syllogisme apparaît comme une instance du schéma

$$\frac{\forall x\,(P(x) \Rightarrow Q(x))\,,\ P(a)}{Q(a)}$$

Ce schéma est valide, comme nous le verrons plus loin.

On peut aussi voir la logique prédicative comme une généralisation de la logique propositionnelle, dans laquelle les conjonctions et les disjonctions peuvent être infinies et se traduisent en quantifications universelles et quantifications existentielles, respectivement.

1.3 Trop simple, la logique ?

On peut se demander à quoi sert la logique en tant que science, puisqu'elle ne recouvre, dans le contexte élémentaire dont nous ne sortirons pas, que des connaissances évidentes. Nous aurons amplement l'occasion de souligner plus loin qu'il est certaines évidences méritant d'être soulignées mais on peut déjà noter une analogie entre l'utilité de la logique et celle de l'arithmétique. Tout d'abord, le fait qu'une règle telle que le Modus Tollens soit aussi élémentaire que la règle arithmétique de commutativité de l'addition ($a + b = b + a$, quels que soient les nombres a et b) n'enlève rien à l'intérêt de ces règles ; en outre, en logique comme en arithmétique existe un effet de dimension qui relativise le caractère élémentaire souvent attribué à certains résultats. En arithmétique, on "produit" $(6 \times 6)/2 = 18$ sans effort, mais il n'en va pas de même pour $345\,234 \times 765\,864 = 264\,402\,292\,176$; on utilisera ici une calculatrice, ou un ordinateur dont la programmation a requis la mise en œuvre de règles arithmétiques bien formalisées. Plus peut-être que la taille des problèmes, c'est leur généralisation qui requiert l'élaboration d'une science. L'égalité

$$1 + 2 + \cdots + 10 = 55$$

présente un intérêt nettement moindre que l'égalité

$$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}\,,$$

parce que la seconde égalité est valable non seulement quand $n = 10$ (auquel cas elle se réduit à la première) mais en fait quand n est n'importe quel entier positif. La première égalité peut se déduire de tables d'addition, alors que pour la seconde une "science" arithmétique est nécessaire. Il en va de même en logique, où l'on formalise des règles générales, telle la classique règle de récurrence :

$$\frac{P(0)\,,\ \forall n\,[P(n) \Rightarrow P(n+1)]}{\forall n\,P(n)}$$

L'effet de dimension est également présent en logique, comme en témoignent deux petites énigmes amusantes.

Enigme "facile".
Les étudiants ayant participé à l'examen de logique diffèrent par le prénom, la nationalité et le sport favori pratiqué par chacun d'eux. On demande de reconstituer le classement et de déterminer qui est le Français et quel est le sport pratiqué par Richard, sur base des indices suivants.

1. *Il y a trois étudiants.*
2. *Michel joue au football.*
3. *Michel est mieux classé que l'Américain.*
4. *Simon est Belge.*
5. *Simon a surclassé le joueur de tennis.*
6. *Le nageur s'est classé premier.*

Enigme "difficile".
Les occupants de maisons alignées diffèrent par la nationalité, la couleur de la maison, la marque de cigarette favorite, la boisson préférée et l'animal familier. On demande de reconstituer la situation, et en particulier d'identifier le propriétaire du zèbre et le buveur d'eau, sur base des indices suivants.

1. *Les numéros des maisons sont 1, 2, 3, 4, 5.*
2. *L'Anglais habite la maison verte.*
3. *L'Espagnol possède un chien.*
4. *On boit du café dans la maison rouge.*
5. *On boit du thé chez l'Ukrainien.*
6. *La maison rouge suit la maison blanche.*
7. *Le fumeur de Old Gold élève des escargots.*
8. *On fume des Gauloises dans la maison jaune.*
9. *On boit du lait au numéro 3.*
10. *Le Norvégien habite au numéro 1.*
11. *Le fumeur de Chesterfield et le propriétaire du renard sont voisins.*
12. *Le fumeur de Gauloises habite à côté du propriétaire du cheval.*
13. *Le fumeur de Lucky Strike boit du jus d'orange.*
14. *Le Japonais fume des Gitanes.*
15. *La maison bleue jouxte celle du Norvégien.*

Formaliser ces énigmes, c'est-à-dire les convertir en formules à analyser, est facile. L'analyse proprement dite est tout aussi facile, mais elle prendra par exemple 30 secondes dans le premier cas et 30 minutes dans le second. La raison en est évidente : les deux énigmes sortent du même moule, dont Lewis Carroll aurait été l'un des premiers à se servir, mais la première est de "dimension" 3 alors que la seconde est de "dimension" 5. D'un point de vue opérationnel, si n est la dimension de l'énigme, une situation possible est consituée de n permutations (indépendantes) de listes de n éléments ; cela peut se faire de $(n!)^n$ manières différentes.[9] Pour la première énigme, la solution est donc à choisir parmi $(3!)^3 = 216$ situations possibles mais pour la seconde, c'est parmi $(5!)^5 = 120^5 = 24\,883\,200\,000$ situations a priori possibles qu'il faudra découvrir la solution . . .

1.4 Logique et mathématique

La logique formelle a été développée notamment par des mathématiciens, pour éclairer divers problèmes délicats survenant en algèbre, en analyse, en géométrie, etc. Il importe de reconnaître d'emblée le statut acquis par la logique mathématique : elle contribue au développement d'autres branches des mathématiques et favorise une meilleure compréhension de celles-ci. Inversement, une certaine maturité mathématique favorise l'apprentissage de la logique. Nous disons bien "une certaine maturité" et non "une vaste culture" : l'étudiant ayant suivi quatre heures de mathématiques par semaine pendant six années d'études secondaires peut parfaitement apprendre la logique formelle.

[9]Rappelons ici que l'exponentielle est la fonction qui itère la multiplication ; x^y est le produit de y facteurs égaux à x et on a, par exemple, $2^5 = 32$. La factorielle de n est le nombre de permutations des nombres de 1 à n. Pour $n = 3$, on a six permutations : (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2) et (3,2,1). Plus généralement, $n!$ est le produit des nombres de 1 à n ; on a par exemple $4! = 24$.

Nous ne donnerons ici qu'un exemple de la symbiose entre logique et mathématique, mais il est capital. Dans n'importe quelle branche des mathématiques, un théorème évoque une catégorie d'objets et affirme que tout objet vérifiant l'hypothèse vérifie aussi la thèse. Ceci est un exemple typique d'évidence qu'il est opportun de souligner. Dans le domaine $\mathbb{Z}$ des entiers relatifs, on a le théorème suivant :

Tout carré est positif.

La paraphrase suivante donne lieu à une traduction immédiate :

Pour tout n, n est un carré implique n est positif.

On peut en effet formaliser l'énoncé en

$$\forall n \, [C(n) \Rightarrow P(n)] \, .$$

Comme souvent en mathématique, on sous-entend une partie de l'information ; dans le cas présent, la formule ne reprend pas (notamment) le fait que n représente un entier relatif et non, par exemple, un nombre complexe. Ce théorème exprime que, des quatre classes d'entiers relatifs que l'on peut a priori former (C-P : carré-positif, C-nP : carré-non-positif, nC-P : non-carré-positif, nC-nP : non-carré–non-positif), la seconde est vide. On a en effet

C-P	$0, 1, 4, \ldots, 100, \ldots$
C-nP	
nC-P	$2, 3, 5, \ldots, 99, 101, \ldots$
nC-nP	$-1, -2, -3, \ldots, -100, \ldots$

Cela illustre la règle logique disant qu'un conditionnel $p \Rightarrow q$ est faux si et seulement si l'antécédent p est vrai et le conséquent q est faux.[10] En particulier, un conditionnel dont l'antécédent est faux est toujours vrai. Par exemple, l'énoncé "si 2+2=5, alors 2+2=6" est vrai, ce qui ne l'empêche pas d'être sans intérêt pratique.

L'apprenti mathématicien éprouve parfois quelque difficulté à absorber la notion de théorème, comme le logicien débutant peut trouver curieuse la règle sémantique relative au conditionnel dit matériel ; nous voyons ici que l'obstacle est le même dans les deux cas.

[10] Le théorème affirme que la seconde des quatre classes est vide ; il n'interdit pas qu'éventuellement une des trois autres classes soit vide aussi.

2 Logique propositionnelle : syntaxe et sémantique

2.1 Introduction

L'objet de la logique propositionnelle est l'étude des propositions et de certaines opérations qui permettent de les combiner. Nous montrons dans cette section d'où proviennent ces objets et ce qu'ils sont.

2.1.1 Généralités sur les propositions

Une *proposition* est une phrase susceptible d'être vraie ou fausse. En français, on parle souvent de "phrase énonciative". Voici quelques exemples de propositions, écrites en langage naturel, en formalisme mathématique … ou dans un mélange des deux.

1. Un plus deux égalent trois.
2. 1+1 = 3.
3. $\pi > e$.
4. Il n'existe que cinq polyèdres réguliers convexes.
5. Il existe une infinité de nombres premiers x tels que $x + 2$ est aussi premier.
6. Le procompsognathus est un deutérostomien anamniote.
7. Il fera beau à Liège le 29 avril de l'an 2021.
8. $x^2 + y^2 = z^2$.
9. Il pleut.
10. Je donne cours de logique le mardi.
11. Je_donne_cours(matière, jour).
12. Un plus deux égalent trois et la terre tourne autour du soleil.
13. Un plus deux égalent trois parce que la terre tourne autour du soleil.
14. Cette phrase est fausse.
15. La phrase suivante est vraie.
16. La phrase précédente est fausse.
17. Ceci n'est pas une phrase.
18. Cette phrase n'est pas une proposition.

On observe immédiatement que la nature propositionnelle d'une phrase n'implique pas la connaissance automatique de la valeur de vérité de cette phrase. Les trois premiers exemples paraissent non problématiques, mais on pourrait hésiter à leur attribuer une valeur de vérité par méconnaissance du français ou de l'arithmétique élémentaire ; on pourrait aussi ignorer ou refuser les conventions habituelles des mathématiciens concernant les constantes numériques importantes, telles $\pi = 3.14159...$ et $e = 2.71828....$ Les exemples 4 à 7 montrent que l'ignorance est une cause excusable et même inévitable de non-attribution. On peut savoir ce qu'est un polyèdre régulier convexe, mais le non-mathématicien ignore généralement leur

nombre. L'exemple 5 a un sens clair, mais il s'agit d'une conjecture de l'arithmétique : les spécialistes pensent qu'elle est vraie, mais n'ont pas réussi à la démontrer. Le sens de la phrase 6 et a fortiori sa valeur de vérité échapperont au non-biologiste. Enfin, faire une prévision météorologique à très long terme est complètement irréaliste.

Les exemples 8 à 11 posent une difficulté d'une autre nature. Les phrases sont claires et élémentaires, mais leur valeur de vérité dépend du contexte. La pertinence de ce contexte peut apparaître explicitement, via des paramètres tels "x", "y", "z", "matière", "jour", ou de manière implicite : qui est "Je" ? Où et quand est prononcée la phrase "Il pleut" ? On ne peut attribuer une valeur de vérité à ces exemples sans connaître leur contexte.

Les exemples 1 à 11 étaient des propositions *atomiques* ; les exemples 12 et 13 sont des propositions *composées* ; les deux composants sont les propositions atomiques "Un plus deux égalent trois" et "La terre tourne autour du soleil". Ces composants sont *connectés* par "et" et par "parce que".

Les exemples 14 à 16 sont des *paradoxes* ; on les comprend, aucune culture et contexte particuliers ne sont nécessaires à leur analyse, mais il est néanmoins impossible de leur attribuer une valeur de vérité sans aboutir à une contradiction (le cauchemar du logicien !). Ce phénomène est lié à l'*autoréférence* : ces phrases parlent d'elles-mêmes. Les exemples 17 et 18 montrent que l'autoréférence n'implique pas toujours le paradoxe : il s'agit bien de deux propositions, toutes deux fausses.

Dans l'approche formelle de la logique propositionnelle, nous voulons nous affranchir de tous les problèmes non liés au raisonnement proprement dit.[11] Un moyen radical d'y parvenir est de restreindre le sens d'une proposition à sa valeur de vérité, exactement comme, en arithmétique, le sens d'une multiplicité est réduit à sa taille. On évoque le nombre "13" par exemple, sans se soucier d'évoquer une multiplicité concrète comportant 13 éléments. En fait, l'arithmétique ne nous apprend rien sur les nombres eux-mêmes, mais a pour objet les relations qui existent entre les nombres, dont l'existence est tout simplement postulée et admise. Le lexique de l'arithmétique se limite donc aux notations identifiant les nombres (suites de chiffres), aux variables représentant les nombres (souvent des lettres, x, y, a, b, etc.) et aux symboles représentant les opérations par lesquelles on peut combiner et comparer les nombres ($+$, $=$, $<$, etc.). L'écriture $x^2 + y^2 = z^2$ est un énoncé de l'arithmétique ; c'est aussi une proposition. Pour attribuer une valeur de vérité à cette proposition, il faut connaître les valeurs (numériques) de x, y et z, mais rien d'autre (inutile, par exemple, de savoir si les nombres en questions représentent des longueurs, ou des vitesses, ou des nombres de pommes contenues dans des paniers).

En arithmétique, on distingue les énoncés *valides*, qui sont toujours vrais, les énoncés *inconsistants*, ou *contradictoires*, qui sont toujours faux, et les énoncés *contingents*, ou *simplement consistants*, qui sont vrais ou faux selon le contexte, c'est-à-dire selon les valeurs que l'on attribue aux variables qu'ils contiennent. Un énoncé valide peut contenir des variables, tel $x^2 + y^2 \geq 2xy$ ou n'évoquer que des constantes, tel $2 + 3 = 5$. Il en va de même pour les énoncés inconsistants ($x^2+y^2 < 2xy$, $2+3 = 6$). En revanche, un énoncé contingent comporte toujours une variable au moins ($x < 3$).

[11] D'après Larousse, la logique est la "science du raisonnement en lui-même, abstraction faite de la matière à laquelle il s'applique et de tout processus psychologique".

La même classification sera adoptée en logique propositionnelle. Les énoncés *true* et $false \Rightarrow p$ sont valides ; seul le second comporte une variable propositionnelle. L'énoncé $p \Rightarrow q$ est contingent, tandis que l'énoncé $true \Rightarrow false$ est contradictoire.

Deux différences essentielles existent entre la logique et l'arithmétique. Tout d'abord, il n'existe que deux valeurs en logique, contre une infinité en arithmétique ; de plus, la notion de proposition existe en arithmétique (les énoncés arithmétiques sont des propositions, au même titre que les énoncés de mécanique des fluides, par exemple), alors que la notion de nombre n'apparaît pas en logique propositionnelle. En fait, la logique "précède" l'arithmétique, car on ne peut pas faire d'arithmétique sans faire, consciemment ou non, de la logique. En contrepartie, l'arithmétique est "plus riche" que la logique ; on pourra identifier le calcul des propositions à un calcul numérique particulier, mais on ne pourra pas identifier le calcul sur les nombres à une logique propositionnelle particulière.

2.1.2 Généralités sur les connecteurs

Les opérateurs combinant les propositions sont appelés *connecteurs*. La logique des propositions est en fait la logique des connecteurs, comme l'arithmétique est plus la science des opérations (addition et multiplication surtout) que celle des nombres proprement dits.

Les opérations de l'arithmétique (au sens large) sont des fonctions dont les arguments (en général, un ou deux) prennent des valeurs numériques ; la valeur du résultat est numérique, ou une valeur de vérité. Voici quelques opérations courantes :

– L'addition : $+ : \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z} : (x, y) \mapsto x + y.$
– Le passage à l'opposé : $- : \mathbb{Z} \longrightarrow \mathbb{Z} : x \mapsto -x.$
– La divisibilité : $| : \mathbb{Z} \times \mathbb{Z}_0 \longrightarrow \{\mathbf{V}, \mathbf{F}\} : (x, y) \mapsto x|y.$
– La primarité : $\mathrm{Pr} : \mathbb{N}_{0,1} \longrightarrow \{\mathbf{V}, \mathbf{F}\} : x \mapsto \mathrm{Pr}(x).$
  Rappelons que, si x est un nombre entier plus grand que 1, Pr(x) est vrai si x est premier, c'est-à-dire n'est divisible que par lui-même et par 1.

Il existe une infinité d'opérations arithmétiques, et le mathématicien s'autorisera à en créer une nouvelle, qu'il nommera et notera de manière appropriée, dès que le besoin s'en fera sentir. Toutefois, quand on étudie l'arithmétique, on se limite généralement à une demi-douzaine d'opérations. On retient, d'une part, celles dont l'intérêt pratique est évident et, d'autre part, celles dont les propriétés sont les plus attrayantes et les plus élégantes. Ces deux critères sont souvent concordants ; de plus, les opérations non retenues comme primitives peuvent souvent se dériver des opérations primitives ... au moyen de la logique. On définit par exemple la divisibilité (ne pas confondre avec la division) à partir de l'égalité et de la multiplication :

$$x|y \ =_{def} \ \exists z\,(x.z = y)\,.$$

On se sert de cette nouvelle opération pour définir la primarité :

$$\mathrm{Pr}(x) \ =_{def} \ x > 1 \,\wedge\, \forall y\,[y|x \,\equiv\, (y = 1 \vee y = x)]\,.$$

En logique propositionnelle aussi, nous aurons des connecteurs fondamentaux et des connecteurs dérivés.

Notons aussi que certaines “opérations” arithmétiques ne sont pas considérées comme telles par les mathématiciens, parce qu’elles dépendent non seulement des nombres eux-mêmes mais aussi de points annexes, par exemple le formalisme utilisé pour les représenter. Ainsi, la “longueur” d’un nombre n’est pas une véritable opération arithmétique, puisqu’elle n’est pas “numérifonctionnelle” :

$$\begin{aligned}
&\ell(13) = 2\,,\\
&\ell(6+7) = 3\,,\\
&\ell(78/6) = 4\,,\\
&\ell(\mathsf{IIOI}) = 4\,,\\
&\ell(\text{treize}) = 6\,,\\
&\ell(\text{thirteen}) = 8\,,\\
&\ell(\text{XIII}) = 4\,.
\end{aligned}$$

En logique aussi, les connecteurs non “vérifonctionnels” seront éliminés.

Les connecteurs propositionnels sont nombreux dans la langue française ; nous en avons rencontré deux exemples :

– Un plus deux égalent trois *et* la terre tourne autour du soleil.
– Un plus deux égalent trois *parce que* la terre tourne autour du soleil.

Le connecteur *et* est vérifonctionnel : la proposition “A et B” sera vraie si et seulement si les propositions “A” et “B” sont toutes deux vraies. En revanche, il n’est pas évident d’établir un éventuel lien de cause à effet entre deux faits, et connaître les valeurs de vérité de “A” et de “B” ne permet généralement pas de connaître la valeur de vérité de “A parce que B”. Les propositions composées suivantes, dont nous supposons les composantes vraies, montrent que le connecteur “parce que” n’est pas vérifonctionnel :

– La voiture dérape *parce que* la route est mouillée.
– La route est mouillée *parce que* la voiture dérape.

Dans un contexte où une voiture a dérapé sur une route mouillée, la première proposition composée semble vraie mais la seconde est quasi certainement fausse. Or, les deux propositions simples contenues dans les propositions composées sont vraies ; cela montre que la valeur de vérité d’une proposition composée avec “parce que” ne dépend pas uniquement des valeurs de vérité des composantes.

Voici quelques exemples d’emploi des connecteurs vérifonctionnels les plus fréquemment utilisés en français.

– J’irai au théâtre *ou bien* j’irai au cinéma.
– Il pleut *ou* il vente.
– *S*’il pleut, *alors* la route est mouillée.
– Le ciel est bleu *et* la neige est blanche.
– Il *n*’est *pas* bête.
– *Si* c’est pile *alors* je gagne *sinon* tu perds !
– Elle réussit *si* elle travaille.
– Elle réussit *seulement si* elle travaille.
– Elle réussit *si et seulement si* elle travaille.

– Elle travaille, *donc* elle réussit.
– [Ils n’ont] *ni* Dieu, *ni* maître !

Dans ces exemples, la valeur de vérité de la proposition composée se déduit aisément de la valeur de vérité des composants ; les connecteurs sont donc bien vérifonctionnels. Dans ce cadre, il est possible de rendre compte de la validité de certains raisonnements, tel le suivant :

1. Tous les hommes sont mortels.
2. Si tous les hommes sont mortels et si Socrate est un homme,
alors Socrate est mortel.
3. Socrate est un homme.
4. Donc, Socrate est mortel.

Ce raisonnement est une *instance*, c’est-à-dire un exemple, du schéma

$$\frac{A\,,\ (A \wedge B) \Rightarrow C\,,\ B}{C}$$

et nous verrons plus loin que toutes les instances de ce schéma (qui comporte trois prémisses et une conclusion) sont valides. On observe cependant que, en bonne logique (informelle) la prémisse

2. Si tous les hommes sont mortels et si Socrate est un homme,
alors Socrate est mortel.

semble redondante, car elle exprime une tautologie, c’est-à-dire une évidence. Comme nous l’avons signalé plus haut, le problème est que la validité du raisonnement

1. Tous les hommes sont mortels.
3. Socrate est un homme.
4. Donc, Socrate est mortel.

ne peut pas être établie dans le cadre du calcul des propositions mais seulement dans celui, plus puissant, du calcul des prédicats. Notons enfin que, ici aussi, des curiosités linguistiques peuvent compliquer l’emploi de la logique en langage naturel. Voici un exemple classique de raisonnement qui, formellement, pourrait sembler valide mais qui, clairement, ne l’est pas.

1. Jacques est un personnage intelligent.
2. Un personnage intelligent a découvert la relativité.
3. Donc, Jacques a découvert la relativité.

En voici un autre :

1. Tout ce qui est rare est cher.
2. Une Rolls-Royce bon marché est rare.
3. Donc, une Rolls-Royce bon marché est chère.

Le calcul des prédicats classique ne pourra pas rendre compte de ces problèmes, qui sont plus du ressort de la linguistique que de la logique.

2.1.3 Les connecteurs vérifonctionnels

Le nombre d'opérations arithmétiques à n arguments est infini et, de plus, il n'est pas possible de donner une table explicite exhaustive pour les opérations arithmétiques, car les opérandes peuvent prendre une infinité de valeurs.[12] Ces limitations n'existent pas en calcul des propositions, puisqu'il n'y a pour les opérandes que deux valeurs possibles. Une table de connecteur sera explicite et exhaustive : on énumère simplement tous les cas possibles. Chaque opérande peut prendre deux valeurs ; pour un opérateur à n arguments, la table comportera donc 2^n lignes. Chaque ligne peut correspondre à un résultat vrai ou à un résultat faux ; il y aura donc 2^{2^n} connecteurs (vérifonctionnels) à n arguments. En particulier, il y a 2 constantes (opérateurs sans argument), 4 connecteurs unaires et 16 connecteurs binaires, dont les tables sont reprises aux figures 2 et 3.

x	$\circ_1$	$\circ_2$	$\circ_3$	$\circ_4$
V	**V**	**V**	**F**	**F**
F	**V**	**F**	**V**	**F**

FIG. 2 – *Les quatre connecteurs unaires.*

x	y	$\circ_1$	$\circ_2$	$\circ_3$	$\circ_4$	$\circ_5$	$\circ_6$	$\circ_7$	$\circ_8$	$\circ_9$	$\circ_{10}$	$\circ_{11}$	$\circ_{12}$	$\circ_{13}$	$\circ_{14}$	$\circ_{15}$	$\circ_{16}$
V	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**	**F**	**F**	**F**	**F**
V	**F**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**	**V**	**V**	**V**	**V**	**F**	**F**	**F**	**F**
F	**V**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**
F	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**	**V**	**F**

FIG. 3 – *Les seize connecteurs binaires.*

Dans la suite, nous n'utiliserons que des connecteurs vérifonctionnels, simplement qualifiés de connecteurs.

2.1.4 Les connecteurs usuels

On voit immédiatement que, des quatre connecteurs unaires, seul le troisième sera vraiment utile ; on l'appelle la *négation*. (Les autres sont les deux constantes et l'identité.) La moitié des connecteurs binaires ont reçu un nom, et un ou plusieurs symboles. Ils sont repris à la figure 4.

Les *tables de vérité* des connecteurs importants, les plus fréquemment utilisés, sont reprises à la figure 5.

Remarque. Les symboles utilisés pour représenter les connecteurs peuvent différer d'un ouvrage à l'autre. Nous avons adopté les notations les plus courantes ; on notera cependant que "$\supset$" est souvent utilisé au lieu de "$\Rightarrow$" ; en Prolog, le symbole "`:-`" est employé à la place de "$\Leftarrow$" et la virgule remplace la conjonction.

[12] La table de multiplication, par exemple, ne concerne que les nombres de 1 à 10 ; des règles supplémentaires sont utilisées pour traiter les cas où les valeurs des opérandes n'appartiennent pas à cet intervalle.

op.	nom	symbole	se lit
$\circ_2$	disjonction	$\vee$	ou
$\circ_3$	conditionnel inverse	$\Leftarrow$	si
$\circ_5$	conditionnel	$\Rightarrow$	si ... alors
$\circ_7$	biconditionnel	$\equiv$	si et seulement si
$\circ_8$	conjonction	$\wedge$	et
$\circ_9$		$\uparrow$	*nand* (en électronique)
$\circ_{10}$	ou exclusif	$\oplus$	*xor*
$\circ_{15}$		$\downarrow$	*nor* (en électronique)

FIG. 4 – *Les connecteurs binaires usuels.*

x	$\neg x$
V	**F**
F	**V**

x	y	$\wedge$	$\vee$	$\equiv$	$\oplus$	$\Rightarrow$
V	**V**	**V**	**V**	**V**	**F**	**V**
V	**F**	**F**	**V**	**F**	**V**	**F**
F	**V**	**F**	**V**	**F**	**V**	**V**
F	**F**	**F**	**F**	**V**	**F**	**V**

FIG. 5 – *Les connecteurs importants.*

Disposer de nombreux connecteurs permet une expression facile et concise des propositions composées, mais rend le formalisme plus complexe, et son étude plus fastidieuse. Avec la négation et un connecteur binaire bien choisi, il est possible de tout exprimer. Supposons par exemple, comme le font souvent les mathématiciens, que les connecteurs "primitifs" sont la négation et le conditionnel. On peut alors introduire les autres connecteurs comme suit :

$$(a \vee b) \ =_{def} \ (\neg a \Rightarrow b)\,,\ (a \wedge b) \ =_{def} \ \neg(a \Rightarrow \neg b)\,, \ldots$$

On montre facilement que $\{\neg, \vee\}$ et $\{\neg, \wedge\}$ constituent aussi des "paires primitives" acceptables, au contraire de $\{\neg, \equiv\}$ et $\{\neg, \oplus\}$. Curieusement, le connecteur binaire $\uparrow$ permet à lui seul de définir tous les autres ; on a par exemple $\neg a \ =_{def} \ (a \uparrow a)$ et $(a \wedge b) \ =_{def} \ ((a \uparrow b) \uparrow (a \uparrow b))$. L'opérateur $\downarrow$ est le seul autre connecteur binaire jouissant de cette propriété.

En français, l'un des rares connecteurs ternaires d'usage courant est "si-alors-sinon". La proposition "si A alors B sinon C" a la valeur de B si A est vrai, et celle de C si A est faux. Ce connecteur permet lui aussi de dériver tous les autres, si on lui adjoint les constantes de base *true* et *false*.[13] Il peut lui-même s'exprimer en termes de connecteurs binaires et de la négation : "si A alors B sinon C" a même valeur de vérité que $((A \wedge B) \vee (\neg A \wedge C))$, ou encore que $((A \Rightarrow B) \wedge (\neg A \Rightarrow C))$.

On peut (on doit !) se demander si ignorer les connecteurs n-aires, avec $n > 2$, ne risque pas d'appauvrir la logique. Pour montrer qu'il n'en est rien, considérons le cas d'un connecteur M à n arguments, avec $n > 2$: la valeur de vérité de $M(p_1, \ldots, p_n)$ dépend (seulement) des

[13] On a par exemple $(a \vee b) \ =_{def}$ [si a alors *true* sinon b].

valeurs de vérité de $p_1, \dots, p_n$. A partir de M, on définit deux autres connecteurs $(n-1)$-aires M^+ et M^-, en posant

$$\begin{aligned} M^+(p_1, \dots, p_{n-1}) &=_{def} M(p_1, \dots, p_{n-1}, \mathit{true})) \, , \\ M^-(p_1, \dots, p_{n-1}) &=_{def} M(p_1, \dots, p_{n-1}, \mathit{false})) \, . \end{aligned}$$

On note immédiatement que les deux formules

$$M(p_1, \dots, p_n)$$

et

$$[(p_n \wedge M(p_1, \dots, p_{n-1}, \mathit{true})) \vee (\neg p_n \wedge M(p_1, \dots, p_{n-1}, \mathit{false}))]$$

sont logiquement équivalentes, c'est-à-dire ont même valeur de vérité, quelles que soient les valeurs de vérité attribuées aux propositions $p_1, \dots, p_n$. La seconde formule, au contraire de la première, ne comporte que des connecteurs à moins de n arguments, ce qui montre que les connecteurs à n arguments ne sont pas indispensables. Plus précisément, on a le résultat suivant :

Théorème. Tout opérateur n-aire ($n > 2$) peut se réduire à une combinaison d'opérateurs binaires et de négations.

Démonstration. La construction donnée plus haut permet d'éliminer un connecteur à 3 arguments en introduisant des connecteurs à 2 arguments, ou encore un connecteur à 27 arguments en introduisant des connecteurs à 26 arguments. En répétant le procédé autant de fois que nécessaire, on arrive à éliminer tout connecteur comportant plus de deux arguments.

Remarque. En logique, les théorèmes affirmant l'existence d'un certain objet se démontrent souvent de façon *constructive* ; la preuve du théorème est une méthode (un algorithme) de construction de l'objet en question. En outre, la preuve se fait souvent par *récurrence*, une notion essentielle que nous allons détailler. Enfin, une fois que l'on sait cela, il suffit de mémoriser une simple ligne pour reconstituer le détail de la preuve. Dans le cas présent, cette ligne peut être

$$M(p_1, \dots, p_n) \equiv [(p_n \wedge M(p_1, \dots, p_{n-1}, \mathit{true})) \vee (\neg p_n \wedge M(p_1, \dots, p_{n-1}, \mathit{false}))] \, .$$

2.2 Digression : la récurrence

2.2.1 Les nombres naturels, de la "définition" à l'axiomatisation

Les nombres naturels 0,1,2,... constituent la base des mathématiques et interviennent dans tous les domaines de la connaissance. Pourtant, il ne semble guère possible de définir l'ensemble $\mathbb{N}$ des entiers naturels ; bien plus, définir un nombre particulier est déjà problématique. On peut dire que le nombre 3 est la caractéristique commune à toutes les collections de trois éléments, mais une telle "définition" pose plus de problèmes qu'elle n'en résout.

Les mathématiciens ont de longue date reconnu ce fait et adoptent généralement le célèbre aphorisme de Leopold Kronecker, l'un des plus brillants mathématiciens du 19ième siècle : en mathématique, Dieu a créé les entiers naturels et l'homme a fait le reste. A défaut de pouvoir définir l'ensemble des entiers naturels, on peut s'efforcer de les caractériser, ou de les résumer, en quelques axiomes les concernant. Voici d'abord trois axiomes "évidents" :

- 0 est un entier naturel.
- Pour tout entier naturel x, il existe un et un seul entier naturel x' appelé son successeur.
- Tout entier naturel y non nul est le successeur x' d'un et un seul entier naturel x appelé son prédécesseur.

Ces trois énoncés traduisent bien notre intuition de la suite des nombres, si on admet que le successeur d'un nombre x est $x + 1$ et que (si $x \neq 0$) son prédécesseur est $x - 1$. Ils ne rendent cependant pas compte de toute notre intuition, en ce sens que l'on peut facilement créer des êtres qui vérifient ces trois axiomes sans s'identifier à nos nombres usuels. Par exemple, on peut ajouter à ceux-ci deux objets particuliers distincts a et b, en décrétant que a et b sont des "entiers naturels", tels que chacun est le prédécesseur et le successeur de l'autre.[14] Nos trois axiomes ne nous permettent pas de chasser ces deux intrus ! On sent aisément où le bât blesse : il doit exister un quatrième axiome, sans doute plus difficile à exprimer, qui permettrait notamment d'éliminer a et b de l'ensemble des naturels. En fait, a et b sont intuitivement à exclure parce qu'ils ne sont pas 0, ni le successeur de 0, ni le successeur de ce successeur, et ainsi de suite. L'axiome "Tout naturel est 0, ou le successeur de 0, ou le successeur du successeur de 0, etc., et rien d'autre" est intuitivement clair mais formellement inacceptable, notamment à cause du "etc.". Une formulation plus acceptable est

- $\mathbb{N}$ est le plus petit ensemble contenant 0 et contenant le successeur de chacun de ses éléments.

2.2.2 La récurrence et sa justification

Du point de vue axiomatique, la dernière proposition est peu satisfaisante, parce que les notions d'ensemble et de plus petit ensemble ne sont a priori pas mieux définies que celle de nombre naturel. Néanmoins, si on accepte ces notions, on doit aussi accepter le principe de récurrence :

> *Si une propriété P est vraie de 0,*
> *et si, chaque fois qu'elle est vraie du naturel n,*
> *elle l'est aussi du naturel successeur n',*
> *alors elle est vraie de tout naturel.*

En effet, supposons que la propriété P vérifie les deux prémisses et notons E_P l'ensemble des naturels pour lesquels P est vraie. L'ensemble E_P est, par définition, inclus dans l'ensemble $\mathbb{N}$ (ce que l'on note $E_P \subset \mathbb{N}$). Mais aussi, vu la première prémisse, E_P contient 0 et, vu la seconde, E_P contient le successeur de chacun de ses éléments ; l'ensemble E_P doit donc inclure $\mathbb{N}$, et donc s'identifier avec lui.

On utilise parfois le raisonnement par l'absurde pour "valider" le principe de récurrence : si la propriété P vérifie les deux prémisses mais pas la conclusion, l'ensemble (noté $\mathbb{N} \setminus E_P$) des naturels pour lesquels la propriété P est fausse n'est pas vide. Soit alors m son plus petit élément ; m ne peut être 0 (première prémisse) ni le successeur d'un naturel (seconde prémisse),[15] d'où une contradiction. On ne peut guère appeler cela une démonstration, parce que l'on tient pour acquis le fait que tout ensemble non vide de nombres naturels admet un minimum, ce qui n'est ni plus ni moins évident que le principe de récurrence lui-même.

[14] On aurait pu introduire aussi un nombre spécial c, qui serait son propre prédécesseur et son propre successeur.
[15] Si $m = x'$ alors, par définition de m, x vérifie P.

Ceci concerne le logicien et le philosophe des sciences pour plusieurs raisons. D'une part, on constate qu'il est difficile de réduire le principe de récurrence à un principe plus "basique", plus évident ; d'autre part, le développement rigoureux de la logique propositionnelle[16] requiert le recours fréquent à ce principe. En fait, la notion de récurrence doit faire partie de la culture générale de tout un chacun, puisqu'elle est inséparable de celle de nombre naturel. Il est d'ailleurs piquant de constater que le premier mathématicien à avoir fait du principe de récurrence un usage conscient et sophistiqué est ... un juriste, Pierre de Fermat.[17]

Concrètement, le principe de récurrence, sous sa forme habituelle

$$\frac{P(0)\,,\ \forall n\,[P(n) \Rightarrow P(n+1)]}{\forall n\,P(n)}$$

est intuitivement évident, car en développant la seconde prémisse et la conclusion, on peut récrire la règle comme suit :

$$\frac{P(0)\,,\ P(0) \Rightarrow P(1)\,,\ P(1) \Rightarrow P(2)\,,\ldots}{P(0)\,,\ P(1)\,,\ P(2)\,,\ldots}$$

On voit que l'on peut obtenir les formules de la conclusion, simplement et de proche en proche, en appliquant la règle de Modus Ponens : $P(0)$ est donné, de $P(0)$ et $P(0) \Rightarrow P(1)$ on déduit $P(1)$, de $P(1)$ et $P(1) \Rightarrow P(2)$ on déduit $P(2)$, etc.

On rencontre aussi la variante dite de récurrence complète :

$$\frac{\forall n\,[(\forall m < n\,P(m)) \Rightarrow P(n)]}{\forall n\,P(n)}$$

Là aussi, une vérification intuitive s'obtient en développant prémisse et conclusion :

$$\frac{P(0)\,,\ P(0) \Rightarrow P(1)\,,\ (P(0) \wedge P(1)) \Rightarrow P(2)\,,\ldots}{P(0)\,,\ P(1)\,,\ P(2)\,,\ldots}$$

Cette vérification intuitive ne constitue pas une véritable démonstration parce que l'on y opère une infinité de déductions élémentaires. Comme nous le verrons plus loin, attribuer aux ensembles infinis des méthodes qui n'ont réellement été définies que pour les ensembles finis expose à des mécomptes.

2.2.3 Utiliser la récurrence

Quelle qu'en soit la raison, chacun se convainc plus ou moins facilement du bien-fondé du principe de récurrence. Un aspect sans doute plus intéressant — et certainement plus utile — du principe de récurrence est son applicabilité à des propriétés particulières. On observe d'abord que l'utilité du principe de récurrence est qu'il permet de démontrer d'un seul coup une infinité d'énoncés élémentaires. Il est fastidieux de démontrer l'égalité

$$0 + 1 + 2 + \cdots + 1000 \ = \ 500\,500\,,$$

[16] et, a fortiori, des logiques plus élaborées.

[17] Il fut Conseiller au Parlement de Toulouse, au 17ième siècle.

mais ce n'est en principe pas difficile. En revanche, démontrer que l'égalité

$$0 + 1 + 2 + \cdots + n \ = \ \frac{n(n+1)}{2}$$

est valable pour tout entier naturel n n'est pas évident. On peut se passer du principe de récurrence en utilisant une astuce :[18] si on appelle $S(n)$ la somme des nombres de 0 à n, on vérifie que $2S(n) = n(n+1)$ par simple inspection du tableau ci-dessous :

$$\begin{array}{c|cccccc} & 0 & 1 & 2 & \ldots & n-1 & n \\ + & n & n-1 & n-2 & \ldots & 1 & 0 \\ \hline = & n & n & n & \ldots & n & n \end{array}$$

Cependant, aucune astuce n'est nécessaire pour appliquer le principe de récurrence. Pour $n = 0$, l'égalité se réduit à $0 = 0(0+1)/2$, ce qui est évident ; la première prémisse du principe est vraie. Pour vérifier la seconde, nous devons *supposer* que l'égalité $S(n) = n(n+1)/2$ est vraie,[19] et *démontrer* que l'égalité $S(n+1) = (n+1)(n+2)/2$ est vraie. On a

$$\begin{array}{rcll} S(n+1) & = & 0 + 1 + \cdots + n + (n+1) & \text{Définition;} \\ & = & (0 + 1 + \cdots + n) \ + \ n+1 & \text{Arithmétique élémentaire;} \\ & = & n(n+1)/2 \ + \ n+1 & \text{Hypothèse de récurrence;} \\ & = & n(n+1)/2 \ + \ 2(n+1)/2 & \text{Arithmétique élémentaire;} \\ & = & (n+2)(n+1)/2 & \text{Arithmétique élémentaire;} \\ & = & (n+1)(n+2)/2 & \text{Arithmétique élémentaire.} \end{array}$$

La méthode de récurrence requiert que l'on connaisse l'énoncé à démontrer avant d'entamer la démonstration ;[20] cela rend parfois très difficile son utilisation, mais ce ne sera pas le cas dans la suite de ce cours. A titre d'exercice, le lecteur pourra vérifier que l'égalité

$$0^2 + 1^2 + 2^2 + \cdots + n^2 \ = \ \frac{n(n+1)(2n+1)}{6}$$

est valable pour tout entier naturel n.[21]

Remarque. Le point de départ de la récurrence n'est pas toujours 0. Par exemple, on peut facilement démontrer par récurrence sur n que, dans tout polygone à n côtés, la somme des angles (exprimée en degrés) vaut $180(n-2)$. Dans ce cas, le point de départ est $n = 3$ (triangle).

Remarque. Dans certains raisonnements mathématiques, la récurrence s'utilise à plusieurs niveaux. Par exemple, après avoir étudié la somme des n premiers entiers naturels, puis celle de leurs carrés et celle de leurs cubes, le mathématicien cherchera naturellement à généraliser ses conclusions et à calculer la somme des $m$ièmes puissances des n premiers entiers naturels ; là aussi, la récurrence (sur m) sera utilisée.

[18] que K.F. Gauss (l'un des plus brillants et précoces mathématiciens de tous les temps) aurait redécouverte à neuf ans.

[19] on peut supposer aussi que $S(k) = k(k+1)/2$ est vraie pour tout $k < n$, c'est inutile ici.

[20] Cette exigence semble minimale mais, en pratique, le mathématicien construit souvent en parallèle la conjecture qu'il veut démontrer et la démonstration elle-même.

[21] Ici, l'astuce de Gauss ne fonctionne pas !

2.2.4 Un usage incorrect de la récurrence

On suppose qu'il existe des billes rouges, des billes vertes et des billes bleues, mais on va "démontrer" (contre toute évidence !) que tout sac de billes est nécessairement unicolore, par récurrence sur le nombre n de billes contenues dans le sac. Pour $n = 1$, la proposition est évidente. Pour $n > 1$, on raisonne comme suit. On ôte une bille a du sac : par hypothèse de récurrence, les $(n-1)$ billes restantes sont d'une seule couleur. On ôte une autre bille b, mais on remet a : les $(n-1)$ billes restantes sont toujours d'une seule couleur, nécessairement la même que précédemment. Donc les n billes sont de la même couleur ! Le lecteur est invité à détecter précisément où se trouve l'erreur, et à spécifier exactement ce que le raisonnement par récurrence a effectivement démontré.

2.2.5 Récurrence non numérique

Fermat a souvent utilisé le principe de récurrence sous une forme particulière, ce qu'il appelait la "méthode de la descente infinie". Pour démontrer qu'une propriété $P(n)$ était vraie pour tout entier n, il supposait qu'elle ne l'était pas pour un certain n_0 et montrait que, si la propriété était fausse pour un certain m, elle l'était encore pour un certain $p < m$ (cette étape pouvait requérir beaucoup d'ingéniosité). Il en déduisait alors l'existence d'une régression infinie, c'est-à-dire d'une suite infinie strictement décroissante $(n_0, n_1, n_2, \ldots)$ d'entiers naturels pour lesquels la propriété serait fausse. Or, une régression infinie ne peut exister : ses éléments formeraient un ensemble d'entiers naturels sans minimum.

Cette variante du principe de récurrence suggère que la récurrence n'est pas applicable uniquement à l'ensemble des entiers naturels, mais à tout ensemble ordonné (totalement ou non) dépourvu de suites infinies décroissantes. L'ensemble des formules de la logique propositionnelle est de ce type. La relation d'ordre est induite par la structure des formules : $\Phi < \Psi$ signifie que Φ est une sous-formule propre[22] de Ψ. En ce sens, les formules minimales sont les propositions élémentaires.

2.3 Syntaxe du calcul des propositions

2.3.1 Les règles de base

Soit $\Pi = \{p, q, r, \ldots\}$, un *lexique propositionnel*, c'est-à-dire un ensemble de symboles arbitraires appelés *propositions atomiques* ou *atomes*. La notation $p \in \Pi$ signifie que p *appartient à* Π, ou *est un élément* de Π. Signalons aussi que l'ensemble vide, celui qui ne contient aucun élément, est noté $\emptyset$.

Définition. Une *formule* du calcul des propositions est une chaîne de symboles générée par la *grammaire*

$$\begin{array}{lll} formula & ::= & p\,,\ \text{pour tout } p \in \Pi \\ formula & ::= & true \mid false \\ formula & ::= & \neg formula \\ formula & ::= & (\textit{formula op formula}) \\ op & ::= & \vee \mid \wedge \mid \Rightarrow \mid \equiv \mid \Leftarrow \end{array}$$

[22] c'est-à-dire un fragment ; cette notion sera formalisée plus loin.

Chaque ligne s'interprète simplement. Par exemple, si nous savons déjà que "$\wedge$" est un opérateur (connecteur) et que "$(p \Rightarrow q)$" et "$\neg r$" sont des formules, la quatrième ligne nous permet de conclure que "$(p \Rightarrow q) \wedge \neg r$" est une formule. On appelle *dérivation* un développement détaillé montrant qu'un assemblage de symboles est une formule.[23] Voici deux exemples de dérivations, montrant que $(p \wedge q)$ et $((p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p))$ sont des formules :

1. *formula*
2. (*formula op formula*)
3. (*formula* $\wedge$ *formula*)
4. ($p \wedge$ *formula*)
5. $(p \wedge q)$

1. $formula$
2. $(formula \equiv formula)$
3. $((formula \Rightarrow formula) \equiv formula)$
4. $((p \Rightarrow formula) \equiv formula)$
5. $((p \Rightarrow q) \equiv formula)$
6. $((p \Rightarrow q) \equiv (formula \Rightarrow formula))$
7. $((p \Rightarrow q) \equiv (\neg formula \Rightarrow formula))$
8. $((p \Rightarrow q) \equiv (\neg q \Rightarrow formula))$
9. $((p \Rightarrow q) \equiv (\neg q \Rightarrow \neg formula))$
10. $((p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p))$

L'ordre des dérivations n'est pas total mais partiel ; il est donc naturel de représenter une dérivation par un arbre. Les arbres de dérivation de la figure 6 montrent l'importance des parenthèses. Deux formules peuvent ne différer que par les positions des parenthèses et avoir des sens très différents.[24]

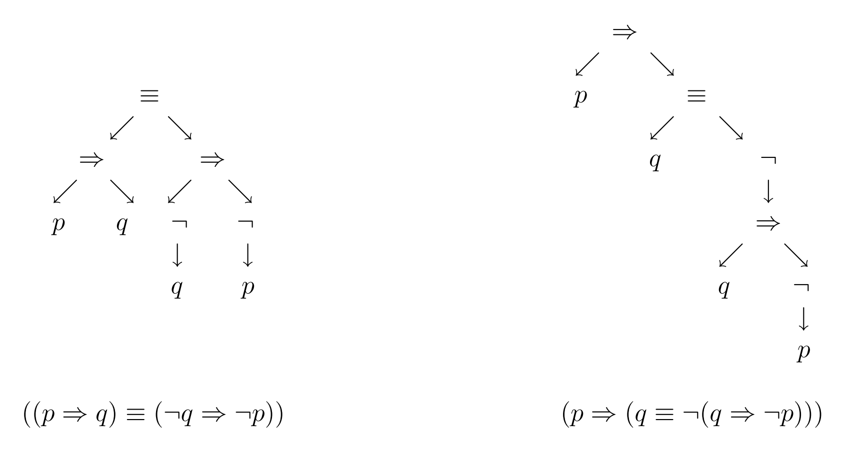


FIG. 6 – Deux arbres de dérivation.

[23] Formellement, cette grammaire comporte deux symboles non terminaux *formula* et *op*. Une formule est donc un mot du langage engendré par la grammaire, dépourvu de symboles non terminaux. C'est le dernier terme d'une dérivation dont le premier terme est le symbole non terminal distingué *formula*.

[24] Le même phénomène se produit en arithmétique ; les expressions arithmétiques $a * (b + c)$ et $(a * b) + c$ ont généralement des valeurs différentes.

2.3.2 Les règles simplificatrices

Les règles simplificatrices de la logique propositionnelle sont analogues à celles de l'arithmétique. Elles sont destinées à abréger l'écriture des formules et à en faciliter la lecture. Voici d'abord deux règles d'un emploi habituel.

- Omission des parenthèses extérieures :
  on écrit $p \wedge q$ au lieu de $(p \wedge q)$, et $q \equiv \neg(q \Rightarrow \neg q)$ au lieu de $(q \equiv \neg(q \Rightarrow \neg q))$.
- Utilisation de l'associativité des opérateurs $\wedge$ et $\vee$:[25]
  on écrit $p \vee q \vee r$ au lieu de $(p \vee q) \vee r$ ou $p \vee (q \vee r)$.

Certains ouvrages utilisent en outre les règles suivantes :

- Groupement à gauche des opérateurs non associatifs :
  on écrit parfois $p \Rightarrow q \Rightarrow r$ au lieu de $(p \Rightarrow q) \Rightarrow r$.
- Priorité des opérateurs :
  en arithmétique, on écrit souvent $a + b * c$ pour $a + (b * c)$; de même on convient par exemple que $p \vee q \wedge r$ équivaut à $p \vee (q \wedge r)$ et non à $(p \vee q) \wedge r$.
  La suite $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftarrow$, $\equiv$ reprend les connecteurs logiques, par ordre décroissant de priorité.

Dans la suite, seules les deux premières règles de simplification seront utilisées. On s'autorisera aussi, pour faciliter la lecture, à remplacer certaines paires de parenthèses par des paires de crochets.

2.3.3 Les notations polonaises

Les logiciens polonais ont proposé des notations permettant de se passer complètement des parenthèses. La notation polonaise directe, ou préfixée, consiste à écrire l'opérateur avant ses opérandes ; la notation polonaise inverse, ou postfixée, consiste à écrire l'opérateur après ses opérandes. La notation habituelle est dite infixée. Changer de notation revient à changer l'ordre de parcours des nœuds dans l'arbre de dérivation. A titre d'exemple, les trois parcours possibles pour les arbres de dérivation de la figure 6 sont représentés à la figure 7.

Notation infixée	$((p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p))$	$(p \Rightarrow (q \equiv \neg(q \Rightarrow \neg p)))$
Notation simplifiée	$(p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$	$p \Rightarrow (q \equiv \neg(q \Rightarrow \neg p))$
Notation préfixée	$\equiv \Rightarrow p\ q \Rightarrow \neg\ q\ \neg\ p$	$\Rightarrow p \equiv q \neg \Rightarrow q \neg p$
Notation postfixée	$p\ q \Rightarrow q\ \neg\ p\ \neg \Rightarrow \equiv$	$p\ q\ q\ p \neg \Rightarrow \neg \equiv \Rightarrow$

FIG. 7 – Les notations polonaises.

Les calculatrices de poche Hewlett-Packard proposent ou imposent l'usage de la notation postfixée, ce qui a contribué à rendre cette notation familière aux non-logiciens.[26]

[25] On verra plus loin comment démontrer cette propriété.

[26] Un bon exercice de programmation consiste à écrire les procédures permettant de passer d'une notation à l'autre.

2.3.4 Formules et sous-formules

La formule A est une *sous-formule* de B si l'arbre syntaxique (l'arbre de dérivation) de A est un sous-arbre de l'arbre syntaxique de B ; la formule A est une *sous-formule propre* de B si A est une sous-formule de B, mais A n'est pas identique à B.
Exemples. $p \Rightarrow q$ est une sous-formule propre de $(p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$; en revanche, $p \Rightarrow q$ est une sous-formule impropre de $p \Rightarrow q$; enfin, $q \equiv \neg q$ n'est pas une sous-formule de $(p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$.

On notera aussi qu'une sous-formule peut avoir plusieurs occurrences dans une formule. Par exemple, la (sous-)formule $\neg p$ a deux occurrences dans le biconditionnel $\neg p \equiv \neg(p \Leftarrow \neg p)$; la (sous-)formule p a trois occurrences dans le conditionnel $p \Rightarrow (p \vee \neg p)$.

2.3.5 Exemples de récurrence non numérique

L'ensemble des formules du calcul des propositions est *inductif* en ce sens qu'il admet un principe de récurrence :

> *Si une propriété P est vraie de toute proposition élémentaire,*
> *et si, chaque fois qu'elle est vraie des formules A et B,*
> *elle l'est aussi des formules $\neg A$, $(A \wedge B)$, $(A \vee B)$, $(A \Rightarrow B)$ et $(A \equiv B)$,*
> *alors elle est vraie de toute formule.*

On appelle souvent *principe d'induction* tout principe de récurrence non numérique.

Ce principe peut être utilisé pour démontrer des propriétés variées. Un exemple simple consiste à démontrer que toute formule (en notation infixée, non simplifiée) comporte un nombre pair de parenthèses. D'une part, c'est vrai pour les propositions élémentaires qui comportent zéro parenthèse. D'autre part, si A comporte α parenthèses, si B comporte β parenthèses et si α et β sont des nombres pairs, alors $\neg A$ comporte α parenthèses (nombre pair) et $(A \wedge B)$, $(A \vee B)$, $(A \Rightarrow B)$ et $(A \equiv B)$ comportent $\alpha + \beta + 2$ parenthèses (nombre pair), d'où la conclusion.

Un propriété plus intéressante affirme que toute formule en notation infixée peut s'écrire en notation préfixée et en notation postfixée. C'est évident pour les propositions élémentaires. D'autre part, si les versions préfixées de A et B sont α et β, alors les versions préfixées de $\neg A$, $(A \wedge B)$, $(A \vee B)$, $(A \Rightarrow B)$ et $(A \equiv B)$ sont respectivement $\neg\alpha$, $\wedge\alpha\beta$, $\vee\alpha\beta$, $\Rightarrow\alpha\beta$ et $\equiv\alpha\beta$. Le principe d'induction permet de conclure. On peut aussi démontrer que la traduction est unique.

2.4 Sémantique du calcul des propositions

2.4.1 Définitions

Une sémantique est une fonction qui donne un sens aux objets d'un langage. Une sémantique est compositionnelle si le sens d'une entité composée au moyen d'un mécanisme donné dépend uniquement du sens des composants. Dans le cadre de la logique des propositions, compositionnel signifie vérifonctionnel. On sait notamment que la formule conjonctive $(A \wedge B)$ est vraie si et seulement si les formules A et B sont toutes les deux vraies. Le caractère non compositionnel des sémantiques des langages naturels est la première cause de leur complexité (et de leur richesse).

Une sémantique compositionnelle est non seulement facile à utiliser mais aussi facile à décrire. Il suffit en fait d'enrichir les règles syntaxiques de construction des formules de règles sémantiques, permettant d'en construire le sens. Aux règles syntaxiques rappelées ci-dessous

$$\begin{array}{lll} formula & ::= & p\,,\ \text{pour tout } p \in \Pi \\ formula & ::= & true \mid false \\ formula & ::= & \neg formula \\ formula & ::= & (\textit{formula op formula}) \\ op & ::= & \vee \mid \wedge \mid \Rightarrow \mid \equiv \mid \Leftarrow \end{array}$$

correspondent des règles sémantiques, permettant d'interpréter, c'est-à-dire de donner un sens à chaque formule.

Une *interprétation* propositionnelle I est basée sur une fonction $v : \Pi \to \{\mathbf{V}, \mathbf{F}\}$ qui attribue une *valeur de vérité* quelconque à chaque atome. L'interprétation I, dont le domaine est l'ensemble des formules construites au moyen du lexique Π, est l'unique fonction respectant les conditions suivantes :

$$\begin{array}{l} I(p) = v(p)\,,\ \text{pour tout } p \in \Pi\,, \\ I(true) = \mathbf{V}\,,\ I(false) = \mathbf{F}\,, \\ I(\neg\varphi) = \mathbf{V} \text{ si } I(\varphi) = \mathbf{F}\,,\ I(\neg\varphi) = \mathbf{F} \text{ si } I(\varphi) = \mathbf{V}\,, \\ I(\varphi\ op\ \psi) = \mathcal{S}(op, I(\varphi), I(\psi))\,, \end{array}$$

la dernière condition est la définition de la fonction sémantique $\mathcal{S}$, résumée à la figure 8.

A	$I(A_1)$	$I(A_2)$	$I(A)$
$A_1 \vee A_2$	**F**	**F**	**F**
$A_1 \vee A_2$	sinon		**V**
$A_1 \wedge A_2$	**V**	**V**	**V**
$A_1 \wedge A_2$	sinon		**F**
$A_1 \Rightarrow A_2$	**V**	**F**	**F**
$A_1 \Rightarrow A_2$	sinon		**V**
$A_1 \Leftarrow A_2$	**F**	**V**	**F**
$A_1 \Leftarrow A_2$	sinon		**V**
$A_1 \equiv A_2$	$I(A_1) = I(A_2)$		**V**
$A_1 \equiv A_2$	$I(A_1) \neq I(A_2)$		**F**

FIG. 8 – Fonction sémantique pour les opérateurs binaires.

Remarque. Conceptuellement, la formule *true* (objet syntaxique) et la valeur de vérité **V** (objet sémantique) sont des objets très différents. L'usage d'une seule notation pour les deux objets est cependant fréquent, et peu gênant en pratique.[27]

Théorème. La fonction d'interprétation v se prolonge en une et une seule interprétation I. C'est une conséquence immédiate de l'unicité de l'arbre syntaxique (arbre de dérivation) d'une formule.

[27]En arithmétique, il est inhabituel de distinguer l'expression arithmétique réduite au seul terme 13 (objet syntaxique) et sa valeur (objet sémantique) qui est le nombre représenté par 13.

Ceci signifie seulement que, si on fixe l'interprétation des propositions élémentaires que contient une formule, on fixe ipso facto l'interprétation de la formule elle-même. La réciproque n'est généralement pas vraie. Si on impose que, par exemple, $p \wedge \neg q$ soit vrai, alors nécessairement p est vrai et q est faux. En revanche, si on impose que $p \wedge \neg q$ soit faux, on ne peut pas déduire avec certitude les valeurs de vérité attachées à p et q.

Remarques. Cet énoncé est effectivement immédiat mais on peut néanmoins le démontrer. Cette démonstration (il en existe plusieurs variantes) met en évidence les caractéristiques des démonstrations en logique : construction de l'objet qu'évoque l'énoncé, raisonnement par récurrence ou par induction. Les règles de la figure 8 permettent de calculer $I(\neg\phi)$ et $I(\phi\ op\ \psi)$ à partir de $I(\phi)$ et $I(\psi)$. On voit donc que si $I(\alpha)$ existe et est unique pour toute formule α comportant (strictement) moins de n connecteurs, alors $I(\alpha)$ existe et est unique pour toute formule α comportant exactement n connecteurs. D'autre part, pour une formule α comportant 0 connecteur, c'est-à-dire une proposition, on a $I(\alpha) = v(\alpha)$ par définition. Cela montre l'existence et l'unicité de l'interprétation I, prolongement sur le domaine des formules de lexique Π de la fonction d'interprétation v (de domaine Π).
On observera aussi que la figure 8 est en fait un *algorithme* (récursif), pour le calcul de $I(\alpha)$. La démonstration ci-dessus est une preuve d'exactitude pour cet algorithme. Notons enfin que, I étant univoquement déterminé à partir de v, il n'est pas indispensable d'utiliser deux notations différentes, ni de distinguer interprétation et fonction d'interprétation. Dans la suite, nous parlerons uniquement d'interprétation, et utiliserons indifféremment I et v pour noter une interprétation quelconque.

La mise en œuvre de la sémantique propositionnelle est élémentaire ; on attribue d'abord une valeur de vérité à toutes les propositions intervenant dans la formule à interpréter, c'est-à-dire à toutes les feuilles de l'arbre syntaxique correspondant à la formule. On "remonte" ensuite dans l'arbre, la valeur d'une sous-formule (correspondant à un nœud interne de l'arbre) dépendant uniquement des valeurs attribuées à ses composants directs. L'arbre syntaxique étant fini, on termine en attribuant une valeur à la formule elle-même, correspondant à la racine de l'arbre.

Exemple. La fonction d'interprétation $v = \{(p, \mathbf{V}), (q, \mathbf{F}), (r, \mathbf{V}), (s, \mathbf{V})\}$ se prolonge en une interprétation I unique sur l'ensemble de toutes les formules basées sur le lexique $\Pi = \{p, q, r, s\}$. Le cas de $(p \vee s) \equiv (s \wedge q)$ est traité à la figure 9.

$$\begin{array}{l} I(p) = \mathbf{V}\,, \\ I(s) = \mathbf{V}\,, \\ I(p \vee s) = \mathbf{V}\,, \\ I(q) = \mathbf{F}\,, \\ I(s \wedge q) = \mathbf{F}\,, \\ I((p \vee s) \equiv (s \wedge q)) = \mathbf{F} \end{array}$$

FIG. 9 – Interprétation d'une formule composée.

Remarque. Le fait que toute fonction d'interprétation v se prolonge en une et une seule interprétation I nous autorise, en pratique, à confondre les deux notions.

Remarque. L'examen exhaustif des interprétations d'une formule composée se fait souvent au moyen d'une *table de vérité* ; cette notion sera approfondie plus loin, mais nous en donnons

déjà un exemple à la figure 10. Cette table montre que la formule $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$ est vraie pour chacune des quatre interprétations possibles.

p	q	$p \Rightarrow q$	$\neg q \Rightarrow \neg p$	$(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$
V	**V**	**V**	**V**	**V**
V	**F**	**F**	**F**	**V**
F	**V**	**V**	**V**	**V**
F	**F**	**V**	**V**	**V**

FIG. 10 – Table de vérité de $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$.

2.4.2 Les connecteurs naturels

Les connecteurs vérifonctionnels utilisés en logique sont des versions simplifiées et idéalisées de leurs homologues naturels. Par exemple, le connecteur "et" de la langue française n'est pas toujours compositionnel, comme le montrent les exemples suivants :

- Il eut peur et il tua l'intrus.
- Il tua l'intrus et il eut peur.

Dans un contexte où les deux composants sont vrais, un procès d'assises par exemple, les deux phrases ont des sens nettement différents. Néanmoins, dans la plupart des cas, le connecteur "et" s'emploie de manière vérifonctionnelle, avec parfois une surcharge de sens, indiquant une nuance temporelle ou causale. Il en va de même des versions naturelles de la négation, de la disjonction et du biconditionnel.

Le connecteur conditionnel pose toutefois un problème. Des phrases telles que

- Si $2 + 2 = 4$, alors la terre tourne autour du soleil.
- Si la terre tourne autour du soleil, alors $2 + 2 = 4$.
- Si $2 + 2 = 5$, alors $2 + 2 = 6$

sont vraies selon les règles sémantiques de la logique propositionnelle,[28] elles pourraient bien être tenues pour fausses (ou "absurdes") par le non-logicien. Il se fait que, la plupart du temps, la notion d'implication n'est pas vérifonctionnelle.

Nous avons déjà noté dans le chapitre introductif que l'implication logique correspondait exactement à l'implication mathématique. Cette implication exprime que le conditionnel reliant l'hypothèse et la thèse est valide, toujours vrai. Il n'empêche que, dans les théorèmes "utiles", le lien entre l'hypothèse (ou la conjonction des hypothèses) et la thèse n'est pas seulement vérifonctionnel ; il s'y ajoute un autre lien, l'existence d'une démonstration permettant de "passer" de l'hypothèse à la thèse. Il n'en est pas moins vrai qu'un énoncé tel que "si $2+2 = 5$ alors $2 + 2 = 6$" est un théorème de l'arithmétique, aussi valide qu'inutile.

On peut aborder le problème autrement. Il n'existe que 16 connecteurs binaires, et donc 16 possibilités de définir une approximation vérifonctionnelle du conditionnel. En outre, certains

[28] On admet que les propositions élémentaires contenues dans ces phrases ont les valeurs de vérité que leur assignent les mathématiques et l'astronomie ...

choix sont d'office exclus. En particulier, on admet aisément que la valeur de vérité d'un conditionnel $(p \Rightarrow q)$ ne peut dépendre du seul antécédent p, ou du seul conséquent q ; on admet de même que les rôles de l'antécédent et du conséquent ne sont pas interchangeables. Si l'on se réfère à la figure 3, les seuls candidats possibles sont $\circ_3$, $\circ_5$, $\circ_{12}$ et $\circ_{14}$. Si on admet en outre que, quand l'antécédent est vrai, le conditionnel a la valeur du conséquent, il ne reste que $\circ_5$, c'est-à-dire le conditionnel tel qu'il a été défini plus haut.

Une comparaison plus poussée entre la logique et l'arithmétique fournira une autre "justification" de la notion de conditionnel (Fig. 11).

2.4.3 Formalisation d'un texte en langage naturel

Comment formaliser le raisonnement ci-dessous en logique propositionnelle ?

> *Si le récit biblique de la création est strictement exact, le soleil n'a pas été créé avant le quatrième jour. Et si le soleil n'a pas été créé avant le quatrième jour, il ne peut avoir été la cause de l'alternance de lumière et d'obscurité pendant les trois premiers jours. Mais soit le sens du mot "jour" dans la bible est différent du sens habituel, soit le soleil a bien été la cause de l'alternance de lumière et d'obscurité pendant les trois premiers jours. Il s'en suit DONC que le récit biblique de la création n'est pas strictement exact ou que le sens du mot "jour" dans la bible est différent du sens habituel.*

Introduisons les propositions suivantes :

- E : le récit biblique de la création est strictement Exact ;
- Q : le soleil a été créé avant le Quatrième jour ;
- A : le soleil a été la cause de l'Alternance de lumière et d'obscurité pendant les trois premiers jours ;
- D : le sens du mot "jour" dans la bible est Différent du sens habituel.

Le raisonnement se formalise en

$$\frac{E \Rightarrow \neg Q\,,\ \neg Q \Rightarrow \neg A\,,\ D \vee A}{\neg E \vee D}$$

On peut vérifier que toute interprétation rendant vraies les trois prémisses rend vraie la conclusion. En effet, si la conclusion est fausse pour l'interprétation v, on a nécessairement $v(E) = \mathbf{V}$ et $v(D) = \mathbf{F}$. Les valeurs de vérité des trois prémisses sont alors

1. $v(E \Rightarrow \neg Q) = v(\neg Q)$;
2. $v(\neg Q \Rightarrow \neg A)$;
3. $v(D \vee A) = v(A)$.

On vérifie immédiatement qu'aucune des quatre manières d'attribuer des valeurs de vérité à Q et A ne permet de rendre simultanément vraies les formules $\neg Q$, $(\neg Q \Rightarrow \neg A)$ et A. En anticipant sur la section suivante, on peut donc affirmer que le raisonnement est correct.

Remarque. Notre analyse ne permet évidemment pas de porter un jugement sur la véracité des trois prémisses. Une simple lecture de la Genèse permet de vérifier la véracité de la première. Pour admettre la seconde, il faut admettre notamment que la cause doit précéder

(temporellement) l'effet. L'analyse de chacune des prémisses et de la conclusion requiert naturellement une bonne connaissance du français. En fait, il est difficile d'inventorier avec précision les connaissances, élémentaires mais nombreuses, nécessaires à la validation de ce raisonnement.

2.4.4 Logique et arithmétique

Si on assimile les valeurs de vérité **F** et **V** aux nombres 0 et 1, respectivement, les connecteurs logiques deviennent les restrictions au domaine $\{0,1\}$ d'opérations arithmétiques. Naturellement, les opérations arithmétiques dont les connecteurs logiques sont des restrictions ne sont pas univoquement déterminées. Par exemple, la fonction identité sur $\{0,1\}$ est la restriction non seulement de la fonction identité sur $\mathbb{N}$ mais aussi, entre autres, de la fonction carré puisque $0^2 = 0$ et $1^2 = 1$. Il est intéressant de considérer que les connecteurs sont les restrictions d'opérations arithmétiques aussi élémentaires et utiles que possible. Nous proposons les rapprochements repris à la figure 11.

$a \equiv b$	$a = b$
$a \Rightarrow b$	$a \leq b$
$a \wedge b$	$\min(a,b)$ ou encore $a * b$
$a \vee b$	$\max(a,b)$
$a \oplus b$	$a \neq b$ ou encore $(a+b) \bmod 2$
$\neg a$	$1 - a$

FIG. 11 – Interprétation arithmétique des connecteurs.

Remarque. L'expression $a \bmod n$, dans laquelle a est un entier et n un entier strictement positif, désigne le reste de la division de a par n, c'est-à-dire l'unique entier $r \in \{0,\ldots,n-1\}$ tel que l'équation $a = nq + r$ admette une solution (nécessairement unique et appelée le quotient).

Le rôle du conditionnel en logique est analogue à celui, crucial, de la relation d'ordre numérique en arithmétique. Comme il n'existe que deux valeurs de vérité, le conditionnel jouit de propriétés particulières. Par exemple, si on considère une formule $A(p)$ comportant une seule occurrence d'une proposition p comme une fonction de p, cette fonction est croissante, si $A(\mathit{false}) \Rightarrow A(\mathit{true})$ est valide (voir paragraphe suivant), ou décroissante (si $A(\mathit{true}) \Rightarrow A(\mathit{false})$ est valide) ou les deux à la fois (si $A(\mathit{false}) \equiv A(\mathit{true})$ est valide). Cette propriété, qui n'est pas vérifiée en arithmétique, permettra de faciliter l'analyse de certaines formules.

2.5 Relation de conséquence logique

Un raisonnement est un mécanisme, ou un algorithme, permettant de dériver une proposition, la *conclusion*, à partir d'un ensemble de propositions données, les *prémisses*. Un raisonnement est *correct*, ou *valide*, si dans tout contexte où les prémisses sont vraies, la conclusion est vraie aussi. On dit alors que la conclusion est *conséquence logique* de l'ensemble des prémisses. Dans cette section, nous abordons le problème central de la logique, qui est de déterminer si une formule est ou n'est pas conséquence logique d'un ensemble donné de formules.

Un problème plus simple est de déterminer si une formule donnée (ou un ensemble de formules) est vrai pour toutes les interprétations possibles, pour au moins une, ou pour aucune. C'est ce problème que nous allons aborder en premier lieu.

2.5.1 Consistance et validité

Consistance et validité d'une formule isolée. Soit A une formule propositionnelle.

- Une interprétation v de A est un *modèle* de A si $v(A) = \mathbf{V}$.
- A est *satisfaisable* ou *consistante* si A a au moins un modèle.
- A est *valide*, ou A est une *tautologie*, si $v(A) = \mathbf{V}$ pour toute interprétation v.
  La notation $\models A$ exprime que A est valide.
- A est *insatisfaisable* ou *inconsistante* si A n'est pas satisfaisable, c'est-à-dire si, pour toute interprétation v, on a $v(A) = \mathbf{F}$.

Remarque. "(In)satisfaisabilité" est le terme propre ; "(in)consistance" est souvent préféré par euphonie.

Théorème. Une formule A est valide si et seulement si sa négation $\neg A$ est insatisfaisable.
Démonstration. Les quatre énoncés suivants sont équivalents.

- A est valide ;
- $v(A) = \mathbf{V}$, pour toute interprétation v ;
- $v(\neg A) = \mathbf{F}$, pour toute interprétation v ;
- $\neg A$ est insatisfaisable.

Consistance et validité d'un ensemble de formules. Soit E un ensemble de formules.

- Le *lexique* Π de E est la réunion des lexiques des éléments de E.
- Une *interprétation* de E est une fonction v de Π dans $\{\mathbf{V},\mathbf{F}\}$; elle admet un prolongement unique permettant d'interpréter toutes les formules dont le lexique est inclus dans Π et, en particulier, toutes les formules de E.
- Une interprétation v de E est un *modèle* de E si elle est un modèle de tous les éléments de E, c'est-à-dire si $v(A) = \mathbf{V}$ pour toute formule $A \in E$.
- E est *satisfaisable* ou *consistant* si E a au moins un modèle.
- E est *insatisfaisable* ou *inconsistant* si E n'est pas satisfaisable, c'est-à-dire si, pour toute interprétation v, on a $v(A) = \mathbf{F}$, pour au moins une formule $A \in E$.

Voici trois conséquences immédiates de ces définitions.

- Toute interprétation est un modèle de l'ensemble vide $\emptyset$.
- Les modèles du singleton $\{A\}$ sont les modèles de la formule A.
- Les modèles de l'ensemble fini $\{A_1,\ldots,A_n\}$ sont les modèles de la conjonction $A_1 \wedge \cdots \wedge A_n$.

Remarques. On peut définir la validité d'un ensemble E comme le fait que toute interprétation est un modèle de E. Cela n'est guère intéressant car un ensemble valide n'est qu'un ensemble de formules valides. La situation est différente en ce qui concerne la consistance. Il est clair que les éléments d'un ensemble consistant sont des formules consistantes, mais un ensemble

de formules consistantes peut être inconsistant ; c'est par exemple le cas de la paire $\{p, \neg p\}$. Une interprétation v d'un ensemble *fini* $E = \{A_1, \ldots, A_n\}$ est un modèle de E si et seulement si c'est un modèle de la formule $A_1 \wedge \cdots \wedge A_n$; c'est pourquoi on parle parfois d'ensemble "conjonctif" de formules. Notons enfin que les notions d'interprétation et de consistance restent pertinentes dans le cas d'un ensemble infini de formules. En revanche, la notion de "conjonction infinie" ou de "formule infinie" n'existe pas dans notre contexte, parce qu'elle correspondrait à un arbre sémantique infini, objet (informatique) difficilement manipulable.

Théorème. Si $E' \subset E$, tout modèle de E est un modèle de E'.
Corollaire. Tout sous-ensemble d'un ensemble consistant est consistant.
Corollaire. Tout sur-ensemble d'un ensemble inconsistant est inconsistant.
Remarque. La notation $E' \subset E$ représente l'énoncé "tout élément de E' est un élément de E".

2.5.2 Conséquence logique, équivalence logique

Une formule A est *conséquence logique* d'un ensemble E de formules si tout modèle de E est un modèle de A. La notation $E \models A$ exprime que A est conséquence logique de E.

Remarque. Si E est valide, et en particulier si E est vide, A est conséquence logique de E si et seulement si A est valide. On peut donc voir la notation

$$\models A$$

comme une abréviation de la notation

$$\emptyset \models A\,.$$

Autrement dit, une formule est valide si et seulement si elle est conséquence logique de l'ensemble vide. On notera aussi qu'une formule est valide si et seulement si elle est conséquence logique de tout ensemble. Dans le même ordre d'idée, la notation

$$E \models \mathit{false}$$

exprime que l'ensemble E est inconsistant. En effet, le seul moyen que tout modèle de E soit un modèle de false est que l'ensemble E n'admette aucun modèle.

Nous introduisons maintenant un résultat, immédiat mais important, permettant de ramener la question "A est-elle conséquence logique de E ?" à la question "$E \cup \{\neg A\}$ est-il inconsistant ?".
Théorème de la déduction (cas fini). Soit A une formule et soit $U = \{A_1, \ldots, A_n\}$ un ensemble fini de formules. Les trois conditions suivantes sont équivalentes :

- A est une conséquence logique de U ; $U \models A$;
- l'ensemble $U \cup \{\neg A\}$ est inconsistant ; $U \cup \{\neg A\} \models \mathit{false}$;
- le conditionnel $(A_1 \wedge \ldots \wedge A_n) \Rightarrow A$ est valide ; $\models (A_1 \wedge \ldots \wedge A_n) \Rightarrow A$.

Théorème de la déduction (cas général). Soit A une formule et soit E un ensemble de formules. Les deux conditions suivantes sont équivalentes :

- A est une conséquence logique de E ; $E \models A$;
- l'ensemble $E \cup \{\neg A\}$ est inconsistant ; $E \cup \{\neg A\} \models \mathit{false}$.

La *théorie* d'un ensemble E de formules est l'ensemble des conséquences logiques de E, soit $\mathcal{T}(E) = \{A : E \models A\}$; les éléments de E sont les *axiomes* ou les *postulats* et les éléments de $\mathcal{T}(E)$ sont les *théorèmes*. Cette notion est surtout employée dans le cadre de la logique des prédicats.

Théorème. Soit E un ensemble de formules et soit U un ensemble de conséquences logiques de E. Les ensembles E et $E \cup U$ admettent exactement les mêmes modèles.
Corollaire. On préserve la consistance d'un ensemble de formules par suppression de formules (quelconques) et aussi par adjonction de conséquences logiques ; on préserve l'inconsistance d'un ensemble par adjonction de formules (quelconques) et aussi par suppression de conséquences logiques (de ce qui n'est pas supprimé !).

Deux formules propositionnelles A_1 et A_2 sont dites *logiquement équivalentes* (ce qui se note $A_1 \leftrightarrow A_2$, ou parfois $A_1 \simeq A_2$) si elles ont les mêmes modèles, c'est-à-dire si $v(A_1) = v(A_2)$ pour toute interprétation v.

En pratique, on peut vérifier simplement l'équivalence logique de deux formules, en passant en revue toutes les interprétations définies sur la réunion des lexiques des deux formules. On établit par exemple

$$p \vee q \leftrightarrow q \vee p$$

en dressant le tableau suivant :

p	q	$v(p \vee q)$	$v(q \vee p)$
V	**V**	**V**	**V**
V	**F**	**V**	**V**
F	**V**	**V**	**V**
F	**F**	**F**	**F**

Remarque. Le mot "équivalence" est employé dans trois cas bien distincts, que nous allons énumérer.

1. Ce mot désigne des objets du *langage formel* qu'est la logique propositionnelle. On peut écrire
   - Le symbole $\equiv$ représente le biconditionnel, parfois nommé aussi l'équivalence.
   - La formule $(p \vee q) \equiv (q \vee p)$ est un biconditionnel valide ;
   - La formule $p \equiv q$ est un biconditionnel contingent ;
   - La formule $p \equiv \neg p$ est un biconditionnel inconsistantt.
2. Ce mot intervient aussi dans le *métalangage*, c'est-à-dire le formalisme, comportant des notations spécifiques, qui permet de parler des objets logiques. On peut écrire
   - Le symbole $\leftrightarrow$ représente la relation d'équivalence logique.
   - L'expression $(p \vee q) \leftrightarrow (q \vee p)$ n'est pas une formule, mais un énoncé du métalangage ; cet énoncé est vrai et exprime que les formules $(p \vee q)$ et $(q \vee p)$ sont logiquement équivalentes.
   - L'énoncé $p \leftrightarrow q$ appartient au métalangage ; il est faux parce que les formules p et q ne sont pas logiquement équivalentes.
3. Enfin, le mot "équivalence" est employé en français, la langue qui nous permet d'écrire ce texte, et d'évoquer les objets du langage et du métalangage. On peut écrire

– Les expressions $\models (A \equiv B)$ et $A \leftrightarrow B$ appartiennent toutes les deux au métalangage ; ce sont des énoncés interchangeables, ou *équivalents*, parce qu'ils sont tous les deux vrais, ou tous les deux faux, selon les formules que les variables (du métalangage) A et B représentent.
– Les énoncés "tout sous-ensemble d'un ensemble consistant est consistant" et "tout sur-ensemble d'un ensemble inconsistant est inconsistant" appartiennent à la langue française (et non au métalangage) ; ils sont *équivalents*, parce qu'ils sont interchangeables ; un raisonnement (informel et élémentaire) permet de déduire un énoncé de l'autre.
– La phrase française "Les énoncés $\models (A \equiv B)$ et $A \leftrightarrow B$ sont équivalents" n'appartient pas au métalangage, mais exprime un fait (vrai) relatif à deux énoncés du métalangage.

L'usage d'un même mot pour désigner des concepts différents se justifie (ou au moins s'explique) par les liens étroits existant entre ces concepts. Ces liens apparaissent par exemple dans le théorème suivant, qui exprime l'équivalence (au sens 3) entre deux énoncés du métalangage.

Pour toutes formules A_1 et A_2, on a $A_1 \leftrightarrow A_2$ si et seulement si on a $\models (A_1 \equiv A_2)$.
Ce théorème, dont la démonstration élémentaire est laissée au lecteur, exprime que deux formules sont logiquement équivalentes (sens 2) si et seulement si le biconditionnel (sens 1) dont elles sont les deux termes est valide.[29]

Selon les formules représentées par A_1 et A_2, les quatre énoncés
– $A_1 \leftrightarrow A_2$.
– $\models (A_1 \equiv A_2)$.
– $\models (A_1 \Rightarrow A_2)$ et $\models (A_2 \Rightarrow A_1)$.
– $\{A_1\} \models A_2$ et $\{A_2\} \models A_1$.

sont, ou bien tous vrais, ou bien tous faux.
Remarques. On écrit souvent $A \models B$ au lieu de $\{A\} \models B$, et $E, A, B \models C$ au lieu de $E \cup \{A, B\} \models C$. De plus, certains auteurs écrivent $v \models A$ au lieu de $v(A) = \mathbf{V}$. Cela vient de ce que l'on assimile parfois l'interprétation v à l'ensemble des formules dont v est un modèle. Mieux vaut éviter cette surcharge de sens pour une notation importante.

2.5.3 Echange et substitution uniforme

En arithmétique et en algèbre, on est souvent amené à remplacer une variable ou une sous-expression d'une expression ou d'une équation donnée par une autre expression. Ce remplacement est intéressant s'il respecte certaines propriétés de l'expression ou de l'équation initiale. Deux cas particuliers sont d'usage fréquent. Tout d'abord, si une expression α contient une sous-expression β égale à une troisième expression γ, on peut remplacer dans α une ou plusieurs occurrences de β par γ sans changer la valeur de α. Supposons par exemple

$$\alpha \;=_{def}\; 2\beta x + 3(\beta{+}\delta)^2(y{-}1)^\beta \text{ et } \beta = \gamma\,.$$

[29]Pour "chicaner" un peu plus, signalons que la locution "si et seulement si" est utilisée, en français, pour exprimer l'équivalence (au sens 3) entre deux énoncés du métalangage … et parfois aussi entre deux énoncés du français, c'est-à-dire entre deux phrases énonciatives quelconques. (Rassurons le lecteur qui nous a suivi jusqu'ici : c'est fini sur ce point !)

On a, entre autres, les égalités suivantes :

$$\alpha = 2\beta x + 3(\gamma{+}\delta)^2(y{-}1)^\beta = 2\beta x + 3(\beta{+}\delta)^2(y{-}1)^\gamma = 2\beta x + 3(\gamma{+}\delta)^2(y{-}1)^\gamma\,.$$

Un autre type de remplacement est souvent employé dans les équations. De l'égalité bien connue

$$(x+y)^2 = x^2 + 2xy + y^2\,,$$

on tire par exemple

$$(ab+3c)^2 = (ab)^2 + 2ab3c + (3c)^2\,.$$

Notons deux différences importantes entre les deux types de remplacements :
– Dans le premier cas on exige l'égalité du terme remplaçant et du terme remplacé, mais pas dans le second.
– Dans le second cas on exige le remplacement uniforme, de toutes les occurrences du terme remplacé, mais pas dans le premier.

Ces deux résultats paraissent évidents mais on doit se méfier, pour au moins trois raisons. La première est que les mathématiques fourmillent de "résultats évidents" … mais faux. La propriété d'associativité de l'addition, souvent résumée par la formule

$$a + (b+c) \;=\; (a+b) + c$$

permet, dans un enchaînement d'additions, de former arbitrairement des résultats intermédiaires, et de calculer indiffemment $(a+(b+c))+d$ ou $(a+b)+(c+d)$, les résultats étant égaux. Cette propriété est valable pour toute somme finie, mais pas pour toute somme infinie (série), comme le montre l'exemple suivant :[30]

$$1 = 1 + [(-1){+}1] + \cdots + [(-1){+}1] + \cdots \stackrel{?}{=} [1{+}(-1)] + \cdots + [1{+}(-1)] + \cdots = 0\,.$$

La deuxième raison est que les résultats susmentionnés, si évidents qu'ils paraissent, deviennent faux dans certains contextes particuliers. Par exemple, on apprend en astronomie que "l'étoile du berger" est en fait une planète, Vénus ; on apprend aussi que les planètes, au contraire des étoiles dites "fixes", tournent autour du soleil. Un imprudent remplacement conduirait à confondre les phrases

Jacques sait que Vénus tourne autour du Soleil.

et

Jacques sait que l'étoile du berger tourne autour du Soleil.

alors que pour beaucoup de gens la première phase est vraie mais pas la seconde. Voici un autre exemple :

L'expression $(x+y)^3$ s'écrit en moins de 10 caractères,
donc l'expression $x^3 + 3x^2y + 3xy^2 + y^3$ s'écrit en moins de 10 caractères.

La troisième raison pour laquelle il n'est pas inutile de démontrer des "résultats évidents" est que les preuves sont parfois aussi instructives que les théorèmes correspondants. En particulier, la validité du premier principe de remplacement évoqué plus haut tient à une propriété essentielle des opérateurs mathématico-logiques ; laquelle ?

[30]Ce fait nous paraît maintenant évident, mais a été dans le passé à l'origine d'erreurs graves.

Lemme de remplacement. La propriété cruciale des opérateurs mathématico-logiques est que la valeur de la forme $f(e_1, \ldots, e_n)$ ne dépend que de l'opérateur f et de la *valeur* des opérandes $e_1, \ldots, e_n$; en particulier, si $e_i = e'_i$, alors $f(e_1, \ldots, e_n) = f(e'_1, \ldots, e'_n)$. On dit que les opérateurs mathématico-logiques sont compositionnels, ou encore dénotationnels ; en logique propositionnelle, on a déjà noté que "compositionnel" signifie "vérifonctionnel". L'opérateur "Je sais" n'est pas vérifonctionnel ; il n'est donc pas étonnant qu'il mette en défaut les principes évoqués plus haut.

Formellement, le lien entre vérifonctionnalité et remplacement s'exprime comme suit.

Lemme (remplacement). Soient A, B, C des formules et v une interprétation. On suppose que A apparaît au moins une fois comme sous-formule de C, et de plus que $v(A) = v(B)$. Si D est le résultat du remplacement d'une occurrence de A par B dans C, alors $v(D) = v(C)$.

Première démonstration du lemme. On considère l'arbre syntaxique relatif à la formule C. Chaque nœud n de l'arbre correspond à une sous-formule φ_n de C ; il peut être étiqueté par la valeur de vérité $v(\varphi_n)$. Le fait que les connecteurs soient vérifonctionnels signifie que la valeur attachée à un nœud intérieur n ne dépend que du connecteur principal de φ_n et des valeurs associées aux descendants directs de n. Le remplacement d'une occurrence de A par B correspond au remplacement d'un sous-arbre par un autre mais, comme $v(A) = v(B)$, ceci ne change ni l'étiquette de la racine du sous-arbre, ni les étiquettes des ancêtres, dont la racine de l'arbre. Cela implique $v(C) = v(D)$.

Remarque. Cette démonstration se résume très bien en un dessin, que nous suggérons au lecteur de tracer. La démonstration qui suit, plus dans le style des mathématiciens, n'implique pas la représentation, même mentale, d'un objet graphique ; on utilise au lieu de cela la notion de profondeur d'une sous-formule dans une formule … ce qui, en fait, revient au même.

Seconde démonstration du lemme. On raisonne par *induction* sur la profondeur d de la sous-formule A dans C, qui correspond à la profondeur de la racine du sous-arbre syntaxique A dans l'arbre C.[31] Soit v une interprétation telle que $v(A) = v(B)$.

– $d = 0$: $A = C$ et $B = D$, donc $v(C) = v(D)$.
– $d > 0$: C est de la forme $\neg C'$ ou $(C'$ *op* $C'')$. Dans le premier cas, A est de profondeur $d - 1$ dans C' et, en nommant D' le résultat du remplacement de A par B dans C', on a (hypothèse inductive) $v(C') = v(D')$; comme $C = \neg C'$ et $D = \neg D'$, on a aussi $v(C) = v(D)$. Dans le second cas, si l'occurrence à remplacer se trouve dans C', on a aussi $v(C') = v(D')$, d'où $v(C) = v(C'$ *op* $C'') = v(D'$ *op* $C'') = v(D)$.[32]

Théorème de l'échange. Le théorème suivant est une conséquence immédiate du lemme de remplacement.

Théorème de l'échange. Soient A et B deux formules ; soient C une formule admettant A comme sous-formule, et D la formule obtenue en remplaçant une ou plusieurs occurrence(s) de A par B dans C. On a

$$(A \equiv B) \models (C \equiv D)\,.$$

Démonstration. Il suffit de montrer que, pour toute interprétation v, si $v(A) = v(B)$, alors $v(C) = v(D)$. Dans le cas particulier où D s'obtient par remplacement d'une seule occurrence de A, on applique le lemme de remplacement. Dans le cas général où n occurrences sont remplacées, il suffit d'appliquer n fois le lemme de remplacement.

Démonstration directe. Considérons une table de vérité pour la formule C. Elle comporte une colonne pour C elle-même, et une colonne pour chacune des sous-formules de C et en particulier une colonne relative à A. Pour obtenir une table de vérité pour D il suffit de

1. Juxtaposer à la colonne relative à A une table relative à B (l'ordre des lignes étant le même).
2. Remplacer les têtes de colonne comportant les occurrences remplacées par la formule résultant du remplacement ; le reste de la colonne n'est pas altéré.
3. Supprimer les colonnes inutiles.

Corollaire. Avec les notations du théorème de l'échange, si A et B sont logiquement équivalents, alors C et D sont logiquement équivalents.[33]

Exemple. Soient les formules

$$\begin{aligned} A &: \ p \Rightarrow q \\ B &: \ \neg p \vee q \\ C &: ((p \Rightarrow q) \wedge r) \vee ((p \Rightarrow q) \Rightarrow r) \\ D &: ((\neg p \vee q) \wedge r) \vee ((p \Rightarrow q) \Rightarrow r) \end{aligned}$$

Une table de vérité relative à C est

p	q	r	$p \Rightarrow q$	$(p \Rightarrow q) \wedge r$	$(p \Rightarrow q) \Rightarrow r$	C
V	**V**	**V**	**V**	**V**	**V**	**V**
V	**V**	**F**	**V**	**F**	**F**	**F**
V	**F**	**V**	**F**	**F**	**V**	**V**
V	**F**	**F**	**F**	**F**	**V**	**V**
F	**V**	**V**	**V**	**V**	**V**	**V**
F	**V**	**F**	**V**	**F**	**F**	**F**
F	**F**	**V**	**V**	**V**	**V**	**V**
F	**F**	**F**	**V**	**F**	**F**	**F**

On introduit les colonnes supplémentaires nécessaires (ici, une seule) :

[31] La notion de profondeur d'une sous-formule peut se définir par induction, sans évoquer la notion d'arbre syntaxique : X est de profondeur 0 dans X ; si X est de profondeur d dans Y, alors X est de profondeur $d + 1$ dans $\neg Y$, dans $Y \Rightarrow Z$, etc. On notera aussi qu'une formule peut contenir plusieurs occurrences d'une sous-formule (cf. § 2.3.4) ; ces occurrences peuvent avoir des profondeurs distinctes.

[32] Ce dernier point utilise explicitement le caractère vérifonctionnel de *op*.

[33] Cela s'écrit, si $\models (A \equiv B)$, alors $\models (C \equiv D)$.

p	q	r	$p \Rightarrow q$	$\neg p \vee q$	$(p \Rightarrow q) \wedge r$	$(p \Rightarrow q) \Rightarrow r$	C
V	**V**	**V**	**V**	**V**	**V**	**V**	**V**
V	**V**	**F**	**V**	**V**	**F**	**F**	**F**
V	**F**	**V**	**F**	**F**	**F**	**V**	**V**
V	**F**	**F**	**F**	**F**	**F**	**V**	**V**
F	**V**	**V**	**V**	**V**	**V**	**V**	**V**
F	**V**	**F**	**V**	**V**	**F**	**F**	**F**
F	**F**	**V**	**V**	**V**	**V**	**V**	**V**
F	**F**	**F**	**V**	**V**	**F**	**F**	**F**

On change les têtes de colonnes concernées par le remplacement (ici, deux), sans changer les colonnes elles-mêmes :

p	q	r	$p \Rightarrow q$	$\neg p \vee q$	$(\neg p \vee q) \wedge r$	$(p \Rightarrow q) \Rightarrow r$	D
V	**V**	**V**	**V**	**V**	**V**	**V**	**V**
V	**V**	**F**	**V**	**V**	**F**	**F**	**F**
V	**F**	**V**	**F**	**F**	**F**	**V**	**V**
V	**F**	**F**	**F**	**F**	**F**	**V**	**V**
F	**V**	**V**	**V**	**V**	**V**	**V**	**V**
F	**V**	**F**	**V**	**V**	**F**	**F**	**F**
F	**F**	**V**	**V**	**V**	**V**	**V**	**V**
F	**F**	**F**	**V**	**V**	**F**	**F**	**F**

Enfin, on supprime les colonnes devenues inutiles (ici, aucune) ; le résultat est une table de vérité pour D.

Corollaire. Avec les notations du théorème de l'échange, si A et B sont logiquement équivalents, alors C et D sont logiquement équivalents.[34]

Exemple. On donne $A =_{def} p$, $B =_{def} \neg\neg p$ et $C =_{def} (p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$. Trois choix sont possibles pour D :

– $D =_{def} (\neg\neg p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$;
– $D =_{def} (p \Rightarrow q) \equiv (\neg q \Rightarrow \neg\neg\neg p)$;
– $D =_{def} (\neg\neg p \Rightarrow q) \equiv (\neg q \Rightarrow \neg\neg\neg p)$.

Comme A et B sont logiquement équivalents, C et D le sont aussi.

Applications du théorème de l'échange. La figure 12 donne une série d'équivalences logiques qui, via le théorème de l'échange, permettent de simplifier des formules. Ces équivalences logiques sont des lois algébriques, analogues aux identités remarquables utilisées en arithmétique et en algèbre.

– On utilise souvent ces lois algébriques, combinées au théorème de l'échange, pour simplifier les formules. Voici un exemple :
$p \wedge (\neg p \vee q) \;\leftrightarrow\; (p \wedge \neg p) \vee (p \wedge q) \;\leftrightarrow\; false \vee (p \wedge q) \;\leftrightarrow\; p \wedge q$

[34] Cela s'écrit, si $\models (A \equiv B)$, alors $\models (C \equiv D)$.

$$
\begin{gathered}
(X \wedge X) \leftrightarrow X \leftrightarrow (X \vee X) \\
(X \wedge Y) \leftrightarrow (Y \wedge X) \\
(X \vee Y) \leftrightarrow (Y \vee X) \\
((X \wedge Y) \wedge Z) \leftrightarrow (X \wedge (Y \wedge Z)) \\
((X \vee Y) \vee Z) \leftrightarrow (X \vee (Y \vee Z)) \\
(X \Rightarrow X) \leftrightarrow true \\
((X \Rightarrow Y) \wedge (Y \Rightarrow X)) \leftrightarrow (X \equiv Y) \\
(((X \Rightarrow Y) \wedge (Y \Rightarrow Z)) \Rightarrow (X \Rightarrow Z)) \leftrightarrow true \\
(X \Rightarrow Y) \leftrightarrow ((X \wedge Y) \equiv X) \\
(X \Rightarrow Y) \leftrightarrow ((X \vee Y) \equiv Y) \\
(X \wedge (Y \vee Z)) \leftrightarrow ((X \wedge Y) \vee (X \wedge Z)) \\
(X \vee (Y \wedge Z)) \leftrightarrow ((X \vee Y) \wedge (X \vee Z)) \\
(X \Rightarrow (Y \Rightarrow Z)) \leftrightarrow ((X \Rightarrow Y) \Rightarrow (X \Rightarrow Z)) \\
(X \vee \neg X) \leftrightarrow true \;;\; (X \wedge \neg X) \leftrightarrow false \\
(X \vee true) \leftrightarrow true \;;\; (X \wedge true) \leftrightarrow X \\
(X \vee false) \leftrightarrow X \;;\; (X \wedge false) \leftrightarrow false \\
\neg\neg X \leftrightarrow X \\
\neg(X \wedge Y) \leftrightarrow (\neg X \vee \neg Y) \\
\neg(X \vee Y) \leftrightarrow (\neg X \wedge \neg Y)
\end{gathered}
$$

FIG. 12 – Quelques équivalences logiques.

– Ces équivalences décrivent des propriétés des connecteurs, telles l'associativité, la commutativité et l'idempotence de $\wedge$, $\vee$, la transitivité du conditionnel et du biconditionnel, etc.
– D'un point de vue sémantique, des formules telles que $p \Rightarrow q$ et $\neg p \vee q$ ne doivent pas être distinguées. L'ensemble des formules construites sur un lexique donné et dans lequel des formules logiquement équivalentes ne "comptent" que pour une seule formule est intéressant à étudier. Cela suggère (aux mathématiciens ...) l'étude de l'ensemble-quotient $\Phi_\Pi / \leftrightarrow$, où Φ_Π désigne l'ensemble des formules basées sur le lexique Π. Cet ensemble-quotient est une *algèbre de Boole* particulière (appelée *algèbre de Lindenbaum*), dont les opérations sont naturellement les connecteurs. Si $\Pi = \{p\}$, l'algèbre correspondante comporte quatre éléments ; on a $A_1 = \{false, p, \neg p, true\}$. Si $\Pi = \{p, q\}$, l'algèbre correspondante comporte seize éléments ; on a $A_2 = \{ false$, $p \wedge q$, $p \wedge \neg q$, $\neg p \wedge q$, $\neg p \wedge \neg q$, p, $\neg p$, q, $\neg q$, $p \oplus q$, $p \equiv q$, $p \vee q$, $p \vee \neg q$, $\neg p \vee q$, $\neg p \vee \neg q$, $true \}$. L'algèbre A_n est isomorphe à $\mathcal{P}(E_n)$, où E_n est un ensemble à 2^n éléments. La négation, la conjonction, la disjonction, le conditionnel et le biconditionnel correspondent respectivement à la complémentation, l'intersection, la réunion, l'inclusion et l'égalité.

Théorème de la substitution uniforme. Soient A_1 et A_2 des formules et B la formule $A_1 \Rightarrow (A_1 \vee A_2)$. La formule B peut être très longue, mais on "voit" immédiatement qu'elle est valide, comme "instance" de la formule valide $p \Rightarrow (p \vee q)$. Si C est une formule, on note

$C[p, q \,/\, A_1, A_2]$ la formule obtenue en remplaçant *toutes* les occurrences des propositions p et q dans C par les formules A_1 et A_2, respectivement. On note aussi $[p, q \,/\, A_1, A_2]$ la fonction (dite "substitution uniforme double") qui à toute formule C associe la formule $C[p, q \,/\, A_1, A_2]$.

Remarque. Les formules $C[p, q \,/\, A_1, A_2]$, $C[p/A_1][q/A_2]$ et $C[q/A_2][p/A_1]$ peuvent être distinctes ; c'est le cas par exemple si C est $p \vee q$ et si A_1 et A_2 sont q et p, respectivement. Dans le cas particulier où aucune proposition p_i n'intervient dans les éléments de l'ensemble $\{A_1, \ldots, A_n\}$, la substitution uniforme n-uple est dite *indépendante* ; on note alors que les formules $C[p_1, \ldots, p_n \,/\, A_1, \ldots, A_n]$ et $C[p_1/A_1] \ldots [p_n/A_n]$ sont nécessairement identiques. Ceci montre que toute substitution uniforme n-uple indépendante est la composée (dans n'importe quel ordre) des n substitutions simples correspondantes. Ce résultat intéresse le programmeur, qui peut remplacer l'affectation n-uple

$$(x_1, \ldots, x_n) := (e_1, \ldots, e_n)$$

par la séquence

$$x_1 := e_1; \ldots ; x_n := e_n$$

à condition que les expressions affectantes ne contiennent pas les variables affectées.

Remarque. Toute substitution uniforme n-uple est la composition de $2n$ substitutions uniformes simples indépendantes.[35]

Lemme de substitution uniforme. Soient $C, A_1, \ldots, A_n$ des formules et $p_1, \ldots, p_n$ des propositions deux à deux distinctes. Si v est une interprétation, définie sur un lexique comportant toute proposition intervenant dans C ou dans l'un des A_i, et telle que $v(p_i) = v(A_i)$ ($i = 1, \ldots, n$), alors on a $v(C[p_1, \ldots, p_n \,/\, A_1, \ldots, A_n]) = v(C)$.

Exemple de substitution uniforme. Soit

$$C \;=_{def}\; p_1 \vee (q \Rightarrow p_2)\,,\; A_1 \;=_{def}\; p_2 \wedge (p_1 \vee r)\,,\; A_2 \;=_{def}\; p_1 \vee q\,.$$

On a alors

$$C[p_1, p_2 \,/\, A_1, A_2] \;=_{def}\; (p_2 \wedge (p_1 \vee r)) \vee (q \Rightarrow (p_1 \vee q))\,.$$

Si on choisit $v \;=_{def}\; \{(p_1, \mathbf{F}), (p_2, \mathbf{V}), (q, \mathbf{V}), (r, \mathbf{F})\}$, on a $v(A_1) = v(p_1) = \mathbf{F}$ et $v(A_2) = v(p_2) = \mathbf{V}$; on a aussi $v([p_1, p_2 \,/\, A_1, A_2]) = v(C) = \mathbf{V}$.

Démonstration du lemme. Dans le cas où la substitution est indépendante, si r_i désigne le nombre d'occurrences de p_i dans C, il suffit d'appliquer $r_1 + \cdots + r_n$ fois le lemme de remplacement (ou n fois le théorème de l'échange). On laisse au lecteur l'extension au cas des substitutions non indépendantes.

Théorème de substitution uniforme. Soient $C, A_1, \ldots, A_n$ des formules et $p_1, \ldots, p_n$ des propositions deux à deux distinctes ; si C est une tautologie, alors $C[p_1, \ldots, p_n/A_1, \ldots, A_n]$ est une tautologie.

Démonstration. On suppose d'abord que la substitution est indépendante ; aucun p_i n'apparaît donc dans $\{A_1, \ldots, A_n\}$, pas plus que dans $C' \;=_{def}\; C[p_1, \ldots, p_n \,/\, A_1, \ldots, A_n]$. Soit v, une interprétation quelconque de C', et w l'*extension* de v obtenue en posant $w(p_i) \;=_{def}\; v(A_i)$.

[35] La démonstration est laissée au lecteur. On pourra utiliser la décomposition d'une affectation n-uple $(x_1, \ldots, x_n) := (e_1, \ldots, e_n)$ en une séquence équivalente de $2n$ affectations simples $t_1 := e_1; \ldots ; t_n := e_n; x_1 := t_1; \ldots ; x_n := t_n$, où les t_i sont n "nouvelles" variables.

Le lemme de substitution uniforme implique $w(C') = w(C)$. Par hypothèse on a $w(C) = \mathbf{V}$ et par construction on a $w(C') = v(C')$. On a donc $v(C') = \mathbf{V}$.

Remarque. Si les p_i intervenaient dans les A_k, la technique pourrait ne pas fonctionner. De $\models p \equiv \neg\neg p$, on ne déduit pas immédiatement que $\models (p \vee r) \equiv \neg\neg(p \vee r)$ car l'interprétation $v : v(p) = \mathbf{F}, v(r) = \mathbf{V}$ telle que $v(A) = v(p \vee r) = \mathbf{V}$ n'admet pas d'extension w telle que $w(p) = \mathbf{V}$. Le remède est simple. De $\models p \equiv \neg\neg p$ on déduit $\models q \equiv \neg\neg q$, d'où on déduit $\models (p \vee r) \equiv \neg\neg(p \vee r)$.

Suite de la démonstration. Si en revanche les p_i interviennent dans $\{A_1, \ldots, A_n\}$, on se donne une famille de nouveaux atomes q_i. Si C est une tautologie, alors $C'' \;=_{def}\; C(p_1/q_1, \ldots, p_n/q_n)$ est une tautologie. D'autre part, C' peut s'écrire $C''[q_1, \ldots, q_n \,/\, A_1, \ldots, A_n]$, où les q_i n'interviennent pas dans les A_k ; C' est donc une tautologie.

Remarque. Où l'exigence d'uniformité de la substitution est-elle utilisée dans cette démonstration ?

Tables de vérité abrégées. On vérifie un énoncé tel que $(p \vee q) \;\leftrightarrow\; (q \vee p)$ au moyen d'une table de vérité (de quatre lignes). Il semble naturel de vérifier un énoncé tel que $(A \vee B) \;\leftrightarrow\; (B \vee A)$ de la même manière, sans recourir à un théorème de substitution uniforme. Cette méthode est acceptable et se justifie comme suit. On peut voir la "pseudo-table" relative à $(A \vee B) \;\leftrightarrow\; (B \vee A)$ comme une abréviation de la table complète, potentiellement très longue, relative à une instance du schéma, telle que

$$[(p \Rightarrow q) \vee r] \;\leftrightarrow\; [r \vee (p \Rightarrow q)]\,.$$

Il est clair que chaque ligne v de la table complète est "représentée" dans la "pseudo-table", par la ligne qui attribue à A la valeur $v(p \Rightarrow q)$ et à B la valeur r. En revanche, certaines lignes de la pseudo-table peuvent ne correspondre à aucune ligne de la table complète. C'est le cas par exemple si A est instancié par une formule valide : les lignes de la pseudo-table concernant les cas où A est faux n'ont pas de correspondant dans la table complète. Cela a la conséquence suivante.

- Si C est une formule valide, alors $C(p_1/A_1, \ldots, p_n/A_n)$ est une formule valide ;
- Si C est une formule inconsistante, alors $C(p_1/A_1, \ldots, p_n/A_n)$ est une formule inconsistante ;
- Si C est une formule simplement consistante, on ne peut rien dire.

Donnons un contre-exemple très simple pour le dernier cas : soit $C \;=_{def}\; p$. On voit que C est simplement consistante, tandis que $C[p/(q \vee \neg q)]$ est valide et que $C[p/(q \wedge \neg q)]$ est inconsistante.

2.6 Quelques théorèmes sémantiques

2.6.1 Interpolation et définissabilité

Introduction. Considérons des fonctions réelles f et g de domaine $\mathbf{R}^2$ et un ensemble $D \subset \mathbf{R}^3$ tels que

$$\forall (x, y, z) \in D : f(x, y) \leq g(x, z)\,.$$

Existe-t-il une fonction *interpolante* $h : \mathbf{R} \longrightarrow \mathbf{R}$ telle que

$$\forall (x,y,z) \in D : [f(x,y) \leq h(x) \leq g(x,z)]?$$

Cela dépend du domaine D. Si par exemple $D = D_1 \times D_2 \times D_3$, deux interpolantes possibles sont

$$x \mapsto \sup_{y \in D_2} f(x,y) \text{ et } x \mapsto \inf_{z \in D_3} g(x,z)$$

Si en revanche on a $D = \{(0,0,0),(0,1,1)\}$, l'hypothèse devient

$$f(0,0) \leq g(0,0) \ \wedge \ f(0,1) \leq g(0,1)$$

et la thèse devient

$$f(0,0) \leq h(0) \leq g(0,0) \ \wedge \ f(0,1) \leq h(0) \leq g(0,1)\,.$$

On voit que, dans ce cas, l'interpolante peut ne pas exister.

Dans $\mathbf{R}^+$, l'équation $x^2 = x + 1$ caractérise un nombre unique (le "nombre d'or"). Autrement dit, si $y^2 = y + 1$ et $z^2 = z + 1$, on a $y = z$, et l'équation définit implicitement le nombre d'or. L'explicitation de ce nombre n'est possible que si on introduit une fonction non rationnelle (la racine carrée) ; on a alors $x = (1 + \sqrt{5})/2$.

La définissabilité et l'interpolation correspondent à la résolution d'équations et d'inéquations. Ces concepts existent aussi en logique, où "$\leq$" et "$=$" deviennent respectivement "$\Rightarrow$" et "$\equiv$".

Théorème d'interpolation de Craig. En logique propositionnelle, l'interpolation doit permettre notamment l'optimisation des circuits digitaux ; une formule comportant n variables propositionnelles distinctes correspond à un circuit digital à n entrées et une sortie. Le plus souvent, un circuit digital n'est pas complètement spécifié et le concepteur peut mettre à profit les degrés de liberté tolérés par la spécification pour obtenir un circuit aussi simple que possible. Dans le cas où la spécification prend la forme d'un intervalle logique, le théorème d'interpolation donne lieu à une technique de simplification.

Théorème. Soient A et B deux formules propositionnelles. Si $\models A \Rightarrow B$, il existe une formule C, ne contenant que des propositions communes à A et B, telle que $\models A \Rightarrow C$ et $\models C \Rightarrow B$.
Démonstration. On raisonne par *induction* sur l'ensemble Π des propositions communes à A et B. Cela signifie que l'on démontre d'abord le théorème dans le cas particulier où l'ensemble Π est vide (cas de base). On suppose ensuite que le théorème est vrai dans le cas d'un ensemble, quelconque mais fixé, ne contenant pas une proposition, elle aussi quelconque mais fixée (cette supposition est l'hypothèse inductive), puis on démontre que le théorème reste vrai dans le cas de cet ensemble augmenté de cette proposition.
Cas de base. Si $\Pi = \emptyset$, $\models A \Rightarrow B$ implique que A est inconsistante (et on choisit $C =_{def} \mathit{false}$) ou que B est valide (et on choisit $C =_{def} \mathit{true}$). Cela se démontre par l'absurde. S'il existait des interprétations u et v (de domaines disjoints) telles que $u(A) = \mathbf{V}$ et $v(B) = \mathbf{F}$, l'interprétation $w =_{def} u \cup v$ serait telle que $w(A \Rightarrow B) = \mathbf{F}$.
Cas inductif. Si $p \in \Pi$, l'hypothèse inductive affirme l'existence d'interpolantes C_T et C_F relatives à $A(p/\mathit{true}), B(p/\mathit{true})$ et à $A(p/\mathit{false}), B(p/\mathit{false})$, respectivement.
On vérifie immédiatement que la formule $(p \wedge C_T) \vee (\neg p \wedge C_F)$ interpole A et B.

Théorème de définissabilité de Beth. *Théorème.* Soit A une formule ne contenant ni q ni r telle que $[A(p/q) \wedge A(p/r)] \Rightarrow (q \equiv r)$ est une tautologie. Il existe une formule B ne contenant pas p, q et r telle que $A \Rightarrow (p \equiv B)$ soit une tautologie.

Remarque. L'hypothèse affirme que A *caractérise* (la valeur de) la proposition p, donc que A définit implicitement p. La thèse affirme qu'une certaine formule B existe, qui définit *explicitement* p. Le théorème affirme donc qu'en logique propositionnelle, une définition implicite peut toujours être explicitée. De ce point de vue, le domaine des formules propositionnelles se distingue de la plupart des domaines mathématiques, dans lesquels l'explicitation des définitions implique souvent l'introduction d'outils nouveaux. Dans une logique plus puissante que la logique propositionnelle, adaptée à l'intelligence artificielle, on peut concevoir, sur base du théorème de Beth, des techniques permettant d'expliciter une information donnée implicitement (résolution d'énigmes par exemple).
Démonstration. On a successivement
$\models [A(p/q) \wedge A(p/r)] \Rightarrow (q \equiv r)]\,,$
$\models [A(p/q) \wedge A(p/r)] \Rightarrow (q \Rightarrow r)]\,,$
$\models [A(p/q) \wedge A(p/r) \wedge q] \Rightarrow r\,,$
$\models [A(p/q) \wedge q] \Rightarrow [A(p/r) \Rightarrow r]\,.$
Soit B une interpolante (théorème de Craig), ne contenant pas p, q, r. On a
$\models [A(p/q) \wedge q] \Rightarrow B$, et donc
$\models A(p/q) \Rightarrow (q \Rightarrow B)$, et par substitution
$\models A(p/r) \Rightarrow (r \Rightarrow B)\,.$
D'autre part, on a
$\models B \Rightarrow [A(p/r) \Rightarrow r]$, d'où
$\models [B \wedge A(p/r)] \Rightarrow r$, d'où
$\models A(p/r) \Rightarrow (B \Rightarrow r)$, et par substitution
$\models A(p/q) \Rightarrow (B \Rightarrow q)\,.$
On en déduit la thèse, sous la forme
$\models A(p/q) \Rightarrow (q \equiv B)$, ou sous la forme
$\models A(p/r) \Rightarrow (r \equiv B)$, ou encore
$\models A \Rightarrow (p \equiv B)\,.$

2.6.2 Théorème de compacité

Préliminaires. La majorité des théorèmes mathématiques évoquent, explicitement ou non, des objets infinis ou des ensembles infinis d'objets. La difficulté vient de ce que les propriétés des objets et ensembles finis ne s'étendent pas systématiquement aux objets et ensembles infinis. Par exemple, l'addition finie est toujours associative et commutative ; l'addition infinie (séries) ne l'est que dans des cas particuliers. Certains ensembles structurés, dits *compacts*, héritent de la plupart des propriétés intéressantes des ensembles finis.

En logique propositionnelle, l'ensemble des interprétations devient infini si l'ensemble des propositions est lui-même infini. La compacité est ici la faculté de ne considérer qu'un sous-ensemble fini (de propositions, d'interprétations, de formules, ...) pour tirer des conclusions portant sur un ensemble infini. Plus concrètement, si E est un ensemble de formules et A une formule, on souhaite que, si A est conséquence logique de E, alors il existe un sous-ensemble fini $E' \subset E$ tel que A soit conséquence logique de E'. D'une manière équivalente,

il faudrait que tout ensemble inconsistant (par exemple $E \cup \{\neg A\}$) admette un sous-ensemble fini inconsistant (par exemple $E' \cup \{\neg A\}$). Ou encore (contraposition), que chaque ensemble *finiment consistant*, c'est-à-dire dont toutes les parties finies sont consistantes, soit lui-même consistant.

Ensembles finiment consistants maximaux. *Définitions.* Un ensemble est *finiment consistant* si tous ses sous-ensembles finis sont consistants.[36] Un ensemble finiment consistant est *maximal* s'il n'admet pas de sur-ensemble finiment consistant.[37]

Théorème. Soient Π un ensemble de propositions et Φ l'ensemble des formules construites sur Π. Un ensemble $E \subset \Phi$ est finiment consistant maximal si et seulement s'il existe une interprétation v sur Π telle que $E = \{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$.
Corollaire. Tout ensemble finiment consistant maximal est consistant et admet un modèle unique.

Démonstration. La preuve montre la correspondance biunivoque entre les interprétations et les ensembles finiment consistants maximaux (pour un lexique fixé).
La condition est suffisante. L'ensemble $E = \{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$ est (finiment) consistant, puisqu'il admet le modèle (unique) v, et est maximal, parce que si $\psi \notin E$, l'ensemble $E \cup \{\psi\}$ contient le sous-ensemble fini inconsistant $\{\neg\psi, \psi\}$.
La condition est nécessaire. On se restreint au cas où le lexique Π est dénombrable, soit $\Pi = \{p_1, p_2, \ldots\}$. Soit E un sous-ensemble f.c. maximal de Φ.

- Pour tout i, E contient exactement un des éléments de la paire $\{p_i, \neg p_i\}$. D'une part, il ne peut contenir les deux éléments, puisque $\{p_i, \neg p_i\}$ est inconsistant. D'autre part, si $p_i \notin E$, l'ensemble $E \cup \{p_i\}$ n'est pas finiment consistant et E admet un sous-ensemble fini E' tel que $E' \cup \{p_i\}$ est inconsistant ; on a alors $E' \models \{\neg p_i\}$. On en déduit que $E \cup \{\neg p_i\}$ est finiment consistant [si $E'' \subset E$, tout modèle de $E'' \cup E'$ est un modèle de $E'' \cup \{\neg p_i\}$] d'où, vu la maximalité, $\neg p_i \in E$.
- Pour tout i, soit ℓ_i l'unique élément de $\{p_i, \neg p_i\}$ appartenant à E ; ces éléments déterminent une interprétation unique, rendant vrais tous les ℓ_i. On note v cette interprétation, dont on va montrer qu'elle est celle requise par l'énoncé.
- On commence par démontrer l'inclusion $E \subset \{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$. Soit $\varphi \in E$ et $\{p_{i_1}, \ldots, p_{i_n}\}$ les propositions intervenant dans φ. Comme E est finiment consistant, son sous-ensemble $\{\ell_{i_1}, \ldots, \ell_{i_n}, \varphi\}$ est consistant, d'où $v(\varphi) = \mathbf{V}$.
- On conclut en observant que, l'ensemble E étant finiment consistant maximal, l'inclusion $E \subset \{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$ entraîne l'égalité $E = \{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$.

Théorème. Tout ensemble finiment consistant est consistant.
Remarque. Il suffit de prouver que tout ensemble finiment consistant est inclus dans un ensemble finiment consistant maximal.
Démonstration. Soit D un ensemble finiment consistant. On pose $E_0 = D$ et, si $n > 0$, $E_n = E_{n-1} \cup \{p_n\}$ si cet ensemble est finiment consistant, $E_n = E_{n-1} \cup \{\neg p_n\}$ sinon.
On démontre par récurrence que tous les E_n sont finiment consistants. C'est trivial pour $n = 0$.

[36] Tout ensemble consistant est donc finiment consistant ; le but de cette section est de montrer que l'inverse est vrai aussi.
[37] On écrira parfois "f.c." et "f.c.max" au lieu de "finiment consistant" et "finiment consistant maximal".

Pour E_n, c'est trivial si $E_{n-1} \cup \{p_n\}$ est finiment consistant. Sinon, il existe un sous-ensemble fini $E' \subset E_{n-1}$ tel que $E' \cup \{p_n\}$ est inconsistant, et donc que $E' \models \neg p_n$. Dans ce cas, $E_n = E_{n-1} \cup \{\neg p_n\}$ est finiment consistant, car pour tout sous-ensemble fini $E'' \subset E_{n-1}$, tout modèle de $E'' \cup E'$ est aussi un modèle de $E'' \cup \{\neg p_n\}$.
On pose $E =_{def} \bigcup_n E_n$. L'intersection $\{p_i, \neg p_i\} \cap E$ contient un élément unique ℓ_i. Ces ℓ_i déterminent une interprétation unique v telle que $v(\varphi) = \mathbf{V}$ pour tout $\varphi \in E$. L'ensemble D, comme l'ensemble E, est donc inclus dans l'ensemble maximal $\{\varphi \in \Phi : v(\varphi) = \mathbf{V}\}$.

Remarque. Le théorème de compacité facilite l'emploi de l'outil logique, mais indique aussi une certaine faiblesse de cet outil. Par exemple, en arithmétique (théorie des nombres entiers), un ensemble infini de formules peut être inconsistant tout en étant finiment consistant. Si z_0 est une constante sur le domaine $\mathbb{Z}$, on pose $E_{z_0} = \{(z > z_0) : z \in \mathbb{Z}\}$. Cet ensemble est inconsistant, puisque pour toute interprétation v, l'entier $v(z_0)$ admet des minorants. Néanmoins, tout sous-ensemble fini de E_{z_0} est consistant. Un tel ensemble s'écrit $\{(z > z_0) : z \in A\}$, où A est un ensemble fini d'entiers. Un modèle v s'obtient en posant $v(z_0) = (\inf A) - 1$. Cela montre simplement que le calcul des propositions ne permet pas d'exprimer toute la théorie arithmétique.

Variante de la démonstration. On peut combiner la construction d'un sur-ensemble maximal et la démonstration du théorème de compacité. Il suffit d'observer que si Π est dénombrable, alors Φ l'est aussi ; en effet, on a $\Phi = \bigcup_n \Phi_n$, où Φ_n est l'ensemble (fini) des formules construites avec le lexique $\{p_1, \ldots, p_n\}$ et comportant au plus n connecteurs. On peut alors considérer une énumération $(\varphi_1, \varphi_2, \ldots)$ de l'ensemble Φ et récrire la démonstration comme suit.

Soit D un ensemble finiment consistant. On pose $E_0 = D$ et, si $n > 0$, $E_n = E_{n-1} \cup \{\varphi_n\}$ si cet ensemble est finiment consistant, $E_n = E_{n-1} \cup \{\neg\varphi_n\}$ sinon.
On démontre par récurrence que tous les E_n sont finiment consistants. C'est trivial pour $n = 0$. Pour E_n, c'est trivial si $E_{n-1} \cup \{\varphi_n\}$ est finiment consistant. Sinon, il existe un sous-ensemble fini $E' \subset E_{n-1}$ tel que $E' \cup \{\varphi_n\}$ est inconsistant, et donc tel que $E' \models \neg\varphi_n$. Dans ce cas, $E_n = E_{n-1} \cup \{\neg\varphi_n\}$ est finiment consistant, car pour tout sous-ensemble fini $E'' \subset E_{n-1}$, tout modèle de $E'' \cup E'$ est aussi un modèle de $E'' \cup \{\neg\varphi_n\}$. On pose $E =_{def} \bigcup_n E_n$. Par construction, pour tout $i > 0$, l'intersection $\{p_i, \neg p_i\} \cap E$ contient un élément unique ℓ_i. Ces éléments déterminent une interprétation v, dont on montre qu'elle est un modèle de E. En effet, soit $\varphi \in E$ et $\{p_{i_1}, \ldots, p_{i_n}\}$ les propositions intervenant dans φ. Soit k le plus petit naturel tel que l'ensemble E_k comporte tous les éléments de $\{\ell_{i_1}, \ldots, \ell_{i_n}, \varphi\}$ (k existe toujours). Comme E_k est finiment consistant, son sous-ensemble $\{\ell_{i_1}, \ldots, \ell_{i_n}, \varphi\}$ est consistant et admet un modèle. Ce modèle ne peut être que v (ou plus exactement la restriction de v au lexique $\{p_{i_1}, \ldots, p_{i_n}\}$), d'où $v(\varphi) = \mathbf{V}$.

Remarquons qu'en arithmétique ce résultat est faux. L'ensemble

$$\{n > 0, n > 1, \ldots, n > 143, \ldots\}$$

est inconsistant, car aucun nombre n'est plus grand que tous les autres, mais tous ses sous-ensembles finis sont consistants.

Pourquoi ce théorème est-il important ? Pour plusieurs raisons, mais nous n'en citons que deux. La première raison est technique. Supposons qu'une formule A soit conséquence logique de l'ensemble infini E de formules. A priori, on pourrait craindre qu'une infinité de formules

de E soient des hypothèses nécessaires pour obtenir la conclusion A. Le théorème de compacité montre que cette crainte n'est pas fondée. En effet, si A est conséquence logique de E, alors $E \cup \{\neg A\}$ est inconsistant et, par le théorème de compacité, il existe un sous-ensemble fini E' de E tel que $E' \cup \{\neg A\}$ soit inconsistant, et donc tel que A soit conséquence logique de E'.

La seconde raison est plus philosophique. L'une des motivations de Frege, dans sa tentative remarquablement réussie de formaliser la logique, était de "réduire" les mathématiques à la logique. Ce "logicisme", dans la lignée du projet leibnizien de "calculus ratiocinator", ne peut réussir que de manière très partielle. Le théorème de compacité (surtout dans le cadre prédicatif, que nous aborderons plus loin) montre que l'arithmétique ne peut se réduire à la logique. Les célèbres résultats d'incomplétude de Gödel montrent que cela a des conséquences importantes.

3 Procédures de décision analytiques

Le théorème de la déduction permet de ramener le problème fondamental de la logique à la détermination de la consistance d'un ensemble de formules, en réduisant la question "la formule A est-elle conséquence logique de l'ensemble de formules E ?" à la question "l'ensemble de formules $E \cup \{\neg A\}$ est-il inconsistant ?". En pratique, on développe surtout des algorithmes de détermination de la consistance d'une formule ou d'un ensemble de formules. Un tel algorithme permet aussi de résoudre le problème de la validité : un ensemble de formules est valide si et seulement si tous ses éléments sont valides[38] et une formule est valide si et seulement si sa négation est inconsistante. Une formule est simplement consistante, ou contingente, si elle n'est ni valide ni inconsistante.

3.1 La méthode des tables de vérité

Le principe de cette méthode est très simple. Toute formule contient un nombre fini d'atomes et admet donc un nombre fini d'interprétations. En conséquence, on peut déterminer la valeur de vérité de la formule pour toutes ses interprétations. On présente souvent le résultat sous forme d'une *table de vérité*, appelée aussi *tableau matriciel*.

On voit immédiatement que la méthode est très inefficace ; si la formule à analyser contient n atomes distincts, elle admet 2^n interprétations. L'algorithme est donc exponentiel (en temps et espace) en fonction du nombre de propositions intervenant dans la formule. Cela signifie que le temps nécessaire à la mise en œuvre de cet algorithme augmente très vite en fonction du nombre de propositions intervenant dans la formule. La figure 13 illustre la méthode des tables de vérité pour trois formules simples.

Le problème de la consistance en logique des propositions est NP-complet. On ne s'attend donc pas à trouver un algorithme polynomial mais, en pratique, on escompte une méthode qui soit exponentielle en pire cas, et non dans tous les cas. La suite de ce chapitre est consacrée à des méthodes qui, le plus souvent, sont nettement meilleures que la méthode des tables de vérité.

On peut améliorer la méthode des tables de vérité en utilisant diverses simplifications. La plus importante consiste à ne pas attendre la fin de la construction de la table pour tirer une conclusion.

Considérons l'exemple de la formule

$$(p \Rightarrow q) \vee (q \Rightarrow r)\,.$$

C'est une disjonction de deux conditionnels. Le premier conditionnel n'est fausx que si p est vrai et q faux, mais dans ce cas le deuxième conditionnel est vrai ; la formule est donc valide.

Nous n'approfondissons pas ici les raffinements que l'on peut apporter à la méthode des tables de vérité, parce qu'il existe d'autres méthodes nettement plus efficaces.[39]

[38] Rappelons ici qu'un ensemble de formules consistantes peut être inconsistant.

[39] La méthode des tables de vérité (avec simplifications) reste intéressante pour résoudre certaines questions théoriques et surtout pour analyser "à la main" des formules très courtes.

p	q	$p \Rightarrow q$	$\neg q \Rightarrow \neg p$	$(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$
V	**V**	**V**	**V**	**V**
V	**F**	**F**	**F**	**V**
F	**V**	**V**	**V**	**V**
F	**F**	**V**	**V**	**V**

Table de vérité de la formule valide $(p \Rightarrow q) \Rightarrow (\neg q \Rightarrow \neg p)$.

p	q	$p \wedge q$
V	**V**	**V**
V	**F**	**F**
F	**V**	**F**
F	**F**	**F**

Table de vérité de la formule simplement consistante $p \wedge q$.

p	q	$p \vee q$	$\neg p$	$\neg q$	$(p \vee q) \wedge \neg p \wedge \neg q$
V	**V**	**V**	**F**	**F**	**F**
V	**F**	**V**	**F**	**V**	**F**
F	**V**	**V**	**V**	**F**	**F**
F	**F**	**F**	**V**	**V**	**F**

Table de vérité de la formule inconsistante $(p \vee q) \wedge \neg p \wedge \neg q$.

FIG. 13 – Application de la méthode des tables de vérité.

3.2 Les tableaux sémantiques

3.2.1 Introduction

La méthode des tables de vérité consiste à passer en revue toutes les interprétations possibles d'une formule φ donnée. La valeur attribuée par une interprétation à la formule φ est calculée en parcourant l'arbre syntaxique de φ du bas vers le haut, des feuilles vers la racine. On utilise le fait que la valeur de vérité d'une formule est déterminée par les valeurs de ses composants.

La méthode des tableaux sémantiques consiste en la recherche systématique d'un modèle d'une formule φ donnée.[40] Le fait d'imposer une valeur de vérité à une formule peut déterminer univoquement la valeur des composants de la formule, ou au contraire laisser plusieurs choix possibles. La recherche systématique d'un modèle conduit à la construction progressive d'une structure arborescente particulière, appelée *tableau sémantique.*[41]

[40] On peut aussi rechercher un antimodèle, une interprétation qui falsifie la formule φ. Il est inutile de considérer explicitement cette variante, puisqu'un modèle de φ est un antimodèle de $\neg\varphi$.

[41] Cette structure est bien un arbre, mais ne doit pas être confondue, d'une part, avec la notion d'arbre syntaxique déjà introduite ni, d'autre part, avec la notion d'arbre sémantique qui sera introduite plus loin.

3.2.2 Technique de construction du tableau

Un exemple introductif. Considérons la formule $\varphi =_{def} (p \Rightarrow q) \wedge \neg(p \Rightarrow r)$, en vue de la recherche systématique d'un modèle. La formule étant une conjonction, tout modèle de φ sera un modèle de ses deux composants (et réciproquement), donc un modèle de l'ensemble $\{p \Rightarrow q\,,\ \neg(p \Rightarrow r)\}$. Le deuxième élément de cet ensemble est la négation d'une implication ; il sera vrai si et seulement si l'antécédent est vrai et le conséquent faux. L'analyse de φ est donc réduite à celle de l'ensemble $\{p \Rightarrow q\,,\ p\,,\ \neg r\}$. Enfin, le premier élément est un conditionnel, que l'on rend vrai en falsifiant l'antécédent ou en vérifiant le conséquent ; il y a là deux possibilités (non exclusives). Les modèles de φ sont donc, d'une part, ceux de l'ensemble $\{\neg p\,,\ p\,,\ \neg r\}$ et, d'autre part, ceux de l'ensemble $\{q\,,\ p\,,\ \neg r\}$.

A ce stade, la décomposition de la formule est achevée parce que les ensembles ne comportent plus que des *littéraux*. Un littéral est un atome (littéral positif) ou la négation d'un atome (littéral négatif). En effet, un ensemble de littéraux est consistant si et seulement s'il ne contient pas simultanément un littéral et son opposé, c'est-à-dire une *paire complémentaire* du type $\{p, \neg p\}$; ceci se détermine par simple inspection. On voit notamment que l'ensemble $\{\neg p\,,\ p\,,\ \neg r\}$ est inconsistant et que l'ensemble $\{q\,,\ p\,,\ \neg r\}$ est consistant ; les modèles de la formule φ sont ceux de ce dernier ensemble.

La méthode des tableaux sémantiques consiste donc à réduire la question (complexe) "la formule A est-elle consistante ?" à la question (triviale) "la famille (finie) $\mathcal{A}$ d'ensembles de littéraux contient-elle un élément consistant ?".

Il est commode d'organiser la recherche sous forme d'un arbre, appelé *tableau sémantique*. Celui correspondant à la formule φ est représenté à la figure 14.

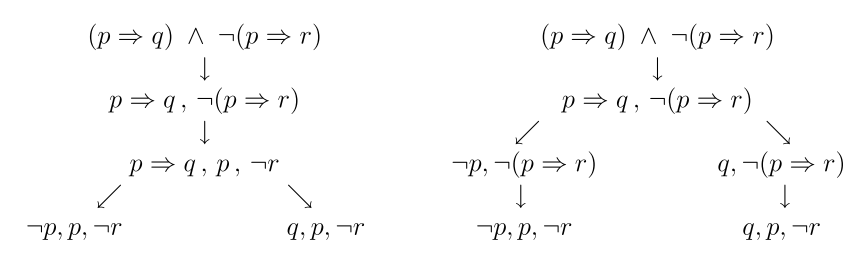


FIG. 14 – Deux tableaux sémantiques.

Une formule peut donner lieu à plusieurs tableaux sémantiques différents suivant l'ordre d'application des règles de construction, mais tous conduisent à la même conclusion (la bonne !) concernant la consistance de la formule. La signification commune des deux tableaux de la figure 14 est

> *Une interprétation est un modèle de la formule* $(p \Rightarrow q) \wedge \neg(p \Rightarrow r)$
> *si et seulement si c'est un modèle de l'un des ensembles* $\{\neg p, p, \neg r\}$ *,* $\{q, p, \neg r\}$.

Construction des tableaux sémantiques. La construction des tableaux sémantiques est basée sur la partition des formules en trois catégories :

– les littéraux ;
– les formules conjonctives ;
– les formules disjonctives.

La formule $\neg(X \Rightarrow Y)$ est conjonctive car elle est équivalente à la conjonction des deux formules (plus simples) X et $\neg Y$. La formule $X \Rightarrow Y$ est disjonctive car elle est équivalente à la disjonction de $\neg X$ et Y. Le connecteur $\equiv$ est exclu ici,[42] et on convient d'assimiler $\neg\neg X$ a X.

La construction est basée sur deux types de règles de décomposition de formules : les règles de *prolongation* (type α) et les règles de *ramification* (type β). Dans le contexte des tableaux sémantiques on utilise les règles α pour les formules conjonctives ; les règles β pour les formules disjonctives.[43] La figure 15 répertorie les formules conjonctives et disjonctives, ainsi que les résultats de l'application à ces formules des règles de prolongation et de ramification, respectivement.[44]

α	α_1	α_2
$A_1 \wedge A_2$	A_1	A_2
$\neg(A_1 \vee A_2)$	$\neg A_1$	$\neg A_2$
$\neg(A_1 \Rightarrow A_2)$	A_1	$\neg A_2$
$\neg(A_1 \Leftarrow A_2)$	$\neg A_1$	A_2

β	β_1	β_2
$B_1 \vee B_2$	B_1	B_2
$\neg(B_1 \wedge B_2)$	$\neg B_1$	$\neg B_2$
$B_1 \Rightarrow B_2$	$\neg B_1$	B_2
$B_1 \Leftarrow B_2$	B_1	$\neg B_2$

FIG. 15 – Les règles de décomposition.

Le processus général de construction d'un tableau sémantique pour une formule donnée φ est décrit à la figure 16. Ce tableau est un arbre dont chaque nœud est étiqueté par un ensemble de formules. Un nœud est *terminal* quand son étiquette ne comporte que des littéraux. Quand la construction du tableau est achevée, toutes les feuilles sont des nœuds terminaux. On convient de marquer un nœud terminal par $\bigcirc$ si l'étiquette est consistante (feuille *ouverte*), et par $\times$ si l'étiquette est inconsistante (feuille *fermée*).

On utilise parfois des *assertions* plutôt que des formules, une assertion étant l'attribution d'une valeur de vérité à une formule. La figure 17 comporte un tableau classique, en notation "formule", et sa variante *signée*, en notation "assertion".

Remarque. Si on adopte cette variante, les liens α-conjonction et β-disjonction ne sont plus valable tels quels ; par exemple, l'assertion $p \wedge q = \mathbf{V}$ est de type α, l'assertion $p \wedge q = \mathbf{F}$ est de type β. En revanche, les liens α-prolongation et β-ramification subsistent. C'est pour éviter ces complications (à tout le moins, ces apparences de complication) que nous n'avons pas directement introduit les tableaux signés, pourtant fréquemment utilisés.

[42] On remplacera donc un biconditionnel $X \equiv Y$ par une formule conjonctive $(X \Rightarrow Y) \wedge (Y \Rightarrow X)$, ou par une formule disjonctive $(X \wedge Y) \vee (\neg X \wedge \neg Y)$, au choix. On élimine de même les formules du type $X \oplus Y$.

[43] Ce point sera nuancé plus loin.

[44] Dans la suite, on notera souvent α une formule conjonctive et β une formule disjonctive ; les composants respectifs seront notés respectivement α_1 et α_2, et β_1 et β_2. Soulignons qu'en général il s'agit de composants sémantiques et non de composants syntaxiques ; par exemple, $\neg p$ n'est pas un composant syntaxique de la formule disjonctive $p \Rightarrow q$.

– *Initialisation.* On crée une racine étiquetée $\{\varphi\}$.
– *Itération.* On sélectionne une feuille non marquée ℓ, d'étiquette $U(\ell)$.
  – Si $U(\ell)$ est un ensemble de littéraux :
    – si $U(\ell)$ contient une paire complémentaire, alors marquer ℓ comme étant fermée ;
    – sinon, marquer ℓ comme étant ouverte.
  – Si $U(\ell)$ n'est pas un ensemble de littéraux, sélectionner une formule dans $U(\ell)$:
    – si c'est une α-formule A, créer un nouveau nœud ℓ', descendant de ℓ, et étiqueter ℓ' avec
$$U(\ell') = (U(\ell) - \{A\}) \cup \{\alpha_1, \alpha_2\};$$
    – si c'est une β-formule B, créer deux nouveaux nœuds ℓ' et ℓ'', descendants de ℓ, et étiqueter ℓ' avec
$$U(\ell') = (U(\ell) - \{B\}) \cup \{\beta_1\}$$
et étiqueter ℓ'' avec
$$U(\ell'') = (U(\ell) - \{B\}) \cup \{\beta_2\}.$$
– *Terminaison.* La construction est achevée quand toutes les feuilles sont marquées '$\times$' ou '$\bigcirc$'.

FIG. 16 – Algorithme de construction d'un tableau sémantique.

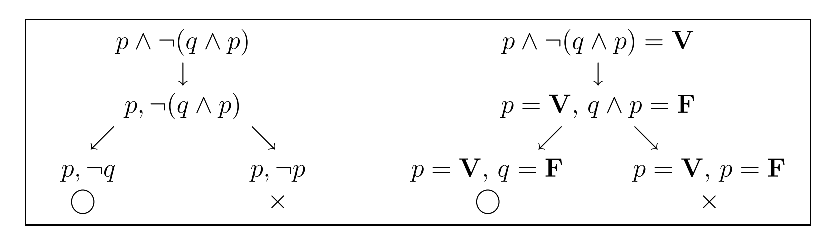
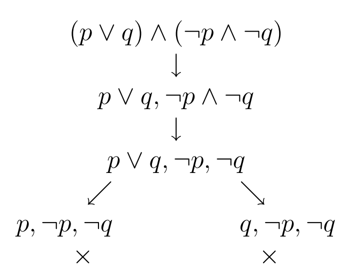


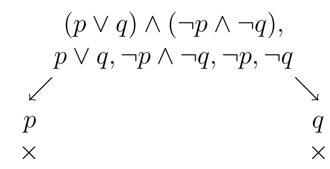
FIG. 17 – Tableau classique et tableau signé.

Programmation de la méthode. La méthode des tableaux sémantiques est facilement programmable. Pour économiser l'espace mémoire, on convient de ne pas créer un nouveau nœud lors de l'application d'une règle α, et de ne pas étiqueter un nœud n avec une formule

qui étiquète déjà un *ancêtre* de n. Avec ces conventions, le tableau



prend la forme plus compacte suivante :



Une autre économie possible est de ne pas étiqueter les nœuds directement avec des sous-formules, mais plutôt avec des pointeurs vers les sous-arbres correspondants dans l'arbre syntaxique. Enfin, certaines branches d'un tableau sémantique sont très semblables. Il est possible d'éviter aussi les redondances de ce type en utilisant non plus une structure d'arbre, mais une structure de graphe sans cycle. Ces raffinements sortent du cadre de ces notes (voir [L2] pour plus de détails).

Terminaison de l'algorithme de construction. *Théorème.* La construction d'un tableau sémantique se termine.
Principe de la démonstration. Chaque étape remplace une formule par une ou deux formules strictement plus simples. Comme la complexité des formules est limitée, on ne peut appliquer qu'un nombre fini d'étapes.
Remarque. On peut, sans restriction essentielle, supposer que le biconditionnel n'est pas employé. Cette restriction simplifie la forme de la mesure W définie dans la démonstration.
Démonstration. Soit $\mathbf{V}$ le tableau d'une formule A, à une étape quelconque de sa construction, soit ℓ une feuille quelconque de $\mathbf{V}$; soient $b(\ell)$ le nombre de connecteurs binaires dans $U(\ell)$ et $n(\ell)$ le nombre de négations dans $U(\ell)$.
On définit $W(\ell) = 2b(\ell) + n(\ell)$. Quelle que soit la règle utilisée, toute étape de la construction crée un nouveau nœud ℓ' ou deux nouveaux nœuds ℓ', ℓ'' tels que $W(\ell') < W(\ell)$ et $W(\ell'') < W(\ell)$. Or, $W(\ell)$ prend ses valeurs dans $\mathbb{N}$; il ne peut donc y avoir de branche infinie dans $\mathbf{V}$.
Remarque. La démonstration peut être étendue au cas où "$\equiv$" et "$\oplus$" apparaissent dans la formule (exercice).

Un tableau *complet* (dont la construction est achevée) est *fermé* si toutes ses feuilles sont fermées. Sinon, il est *ouvert*.

3.2.3 Propriétés de la méthode des tableaux sémantiques

La méthode des tableaux sémantiques est le plus souvent utilisée pour montrer la validité d'une formule (ou l'inconsistance de sa négation), ou encore pour montrer l'inconsistance d'un ensemble fini de formules. On souhaite naturellement que la méthode donne uniquement des résultats corrects ; elle ne peut conclure, par exemple, à l'inconsistance d'une formule, que si la formule est effectivement inconsistante. Cette propriété est *l'adéquation* de la méthode. D'autre part, on souhaite que, si une formule est inconsistante, la méthode mette ce fait en évidence ; c'est la propriété de *complétude*. En résumé, une méthode est adéquate si elle est correcte ; elle est complète si elle est assez puissante.

Dans le cas présent, prouver l'adéquation revient à prouver l'un des énoncés suivants :
- si $T(A)$ est fermé, alors A est inconsistante ;
- si $T(\neg B)$ est fermé, alors B est valide ;
- si A est consistante, alors $T(A)$ est ouvert ;
- si B n'est pas valide, alors $T(\neg B)$ est ouvert.

Prouver la complétude revient à prouver la réciproque, c'est-à-dire l'un des énoncés suivants :
- si A est inconsistante, alors $T(A)$ est fermé ;
- si B est valide, alors $T(\neg B)$ est fermé ;
- si $T(A)$ est ouvert, alors A est consistante ;
- si $T(\neg B)$ est ouvert, alors B n'est pas valide.

3.2.4 Digression : le mouvement et le changement

Le mouvement est impossible, affirmait Zénon d'Elée, puisque pour aller d'un point à un autre, il faut d'abord passer par le point médian, puis, pour aller de ce point médian au point destination, il faut passer par le milieu de ces deux points, etc. Dans la mesure ou le mouvement impliquerait le passage par un nombre infini de points, il ne peut avoir lieu. Un autre argument permettait à Zénon d'affirmer qu'Achille ne rattraperait jamais la Tortue si celle-ci avait une certaine avance, car il faudrait d'abord réduire cette avance de moitié, puis réduire l'avance résiduelle de moitié, et ainsi de suite.

Ces paradoxes anciens ont une conséquence très moderne : puisque ce qui bouge, ou ce qui change, semble moins accessible à l'analyse que ce qui est immobile et immuable, on doit s'efforcer de ramener le premier cas au second. Un exemple classique est l'énigme de la mouche et des deux trains :

> *Deux trains distants de 100 km se dirigent l'un vers l'autre. Ils roulent chacun à 50 km/h. Une mouche quitte le premier train et se dirige vers le second à 100 km/h. Quand elle l'a rejoint, elle fait demi-tour ; chaque fois qu'elle rejoint l'un des trains, elle fait demi-tour. Quelle distance totale aura-t-elle parcouru lorsque les deux trains se croiseront ?*

Si on "suit le mouvement", l'analyse est délicate. La mouche change de direction de plus en plus vite . . . En revanche, il est clair que la mouche vole pendant une heure, à l'allure constante et immuable de 100 km/h : elle aura donc parcouru exactement 100 km.

Tout ceci s'applique à la physique. Pour reprendre l'exemple déjà évoqué des lois de Kepler, il n'est a priori pas facile de prévoir le mouvement d'une planète autour du soleil, puisque la

direction et la vitesse du mouvement changent en permanence. Le génie de Kepler a été de détecter ce qui ne changeait pas : la ligne imaginaire joignant la planète au soleil balaie des surfaces égales en des temps égaux. Cette caractéristique “invariante” du mouvement facilite grandement la prévision de la position des planètes.

Ceci s’applique aussi à la logique et aux procédures de décision. Appliquer une procédure de décision, c’est construire un objet changeant, difficile à analyser, sauf si on se concentre sur ce qui ne change pas. Nous allons appliquer ce principe au prochain paragraphe.

Une propriété particulière de certains processus dynamiques est leur *terminaison*. Certains processus se stabilisent après un certain délai, d’autres ne s’arrêtent jamais. Considérons d’abord un exemple classique. Des nénuphars envahissent un étang. Chaque jour, la moitié de la surface restante est recouverte. Au bout de combien de jours l’étang est-il entièrement recouvert ? Ce problème peut se mathématiser comme suit : dans la suite numérique

$$1, 1/2, 1/4, 1/8, \ldots$$

quel est le rang du premier terme nul ? La réponse est évidente ; aucun terme n’est nul. L’étang ne sera donc jamais entièrement recouvert. Un autre exemple tout aussi évident est le suivant. Un marchand de billes vend au moins une bille par jour. Combien de jours durera son stock ? En termes mathématiques, considérons la suite

$$x_0, x_1, x_2, \ldots$$

telle que chaque terme est un nombre entier naturel strictement inférieur à son prédécesseur. Quelle est la longueur de la suite ? La réponse est évidente : cette longueur est au maximum $x_0 + 1$, si x_i est $x_0 - i$, pour tout i compris entre 0 et x_0.

Peut-on, d’une manière analogue, démontrer que le processus de construction d’un tableau sémantique se termine toujours ? L’idée de démonstration mentionnée plus haut convient, mais on pourrait faire cette objection : à chaque étape, les formules deviennent plus simples, mais aussi plus nombreuses ; le second fait ne compense-t-il pas le premier, et la diminution de complexité est-elle réelle ? Pour répondre avec rigueur à cette question, il faut donner une définition précise de la notion de complexité d’une formule ou d’un ensemble de formules. Plusieurs possibilités existent, par exemple : la complexité d’une formule est le nombre de négations qu’elle contient, plus deux fois le nombre d’opérateurs binaires[45] ; de plus, la complexité d’un ensemble de formules est la somme des complexités des éléments. On voit immédiatement que la complexité est un entier naturel, et que la complexité d’un nœud fils est moindre que celle du nœud père. La “profondeur” du tableau sémantique relatif à une formule de complexité n est donc au plus n ; le nombre total de feuilles est donc au plus 2^n et le nombre total de noeud au plus $2^{n+1} - 1$.

Une façon plus abstraite de présenter cet argument de terminaison est la suivante. On dispose d’un multi-ensemble E_0 de nombres entiers naturels.[46] On construit une suite $(E_0, E_1, E_2, \ldots)$ dans laquelle chaque multi-ensemble s’obtient en remplaçant un élément non nul par un ou deux éléments plus petits. La suite est-elle toujours finie ? Le raisonnement fait plus haut montre que oui. De manière

[45] Conjonction, disjonction, conditionnel et conditionnel inverse ; nous excluons les autres opérateurs.

[46] Un multi-ensemble est un ensemble dans lequel on tient compte des répétitions (mais pas de l’ordre) ; $\{0, 0, 1\}$ et $\{0, 1\}$ sont le même ensemble, mais deux multi-ensembles distincts ; en revanche, $\{0, 0, 1\}$ et $\{0, 1, 0\}$ sont le même multi-ensemble.

plus étonnante, la réponse reste positive si chaque multi-ensemble s’obtient en remplaçant un élément non nul par *un nombre quelconque* d’éléments plus petits.

3.2.5 Adéquation et complétude de la méthode des tableaux sémantiques

L’algorithme de la figure 16 concerne la construction d’un tableau sémantique, c’est-à-dire d’un arbre. Chaque nœud de l’arbre a zéro, un ou deux fils. Les feuilles n’ont pas de fils, les nœuds internes en ont au moins un. La construction consiste en la sélection d’une “feuille provisoire” et la création d’un ou deux fils pour ce nœud qui devient donc un nœud interne. Pendant la construction, l’arbre est un objet changeant puisqu’il grandit. Pour déterminer ses propriétés, il convient d’abord de s’attacher aux caractéristiques non altérées par les changements successifs, qui se résument à deux choses. D’une part, une feuille provisoire comportant α est remplacée par une feuille provisoire comportant α_1 et α_2 ou, d’autre part, une feuille provisoire comportant β est remplacée par deux feuilles provisoires comportant respectivement β_1 et β_2. Or, une interprétation est un modèle de α si et seulement si elle est un modèle de α_1 et de α_2 ; elle est un modèle de β si et seulement si elle est un modèle de β_1 ou de β_2.

Soit $\mathcal{T}_n$ l’arbre à la $n$ième étape de sa construction. On note $\mathcal{M}(\mathcal{T}_n)$ l’ensemble des interprétations qui rendent vraie l’étiquette d’au moins une feuille provisoire de $\mathcal{T}_n$.[47] On constate que $\mathcal{T}_n$ croît avec n, mais $\mathcal{M}(\mathcal{T}_n)$ reste inchangé. Plus précisément, nous venons de voir que la $(n+1)$ième étape, qu’elle soit “de type α” ou “de type β”, n’altère pas l’ensemble des modèles ; on a donc

$$\mathcal{M}(\mathcal{T}_{n+1}) = \mathcal{M}(\mathcal{T}_n)\,,$$

quel que soit l’entier naturel n, et donc

$$\mathcal{M}(\mathcal{T}_n) = \mathcal{M}(\mathcal{T}_0)\,,$$

pour tout n. La valeur commune de ces ensembles d’interprétations est donc l’ensemble des modèles de la formule φ qui étiquète la racine de l’arbre.

Les propriétés annoncées au paragraphe 3.2.3 sont maintenant évidentes. Supposons que la construction du tableau se fasse en N étapes. On a donc $\mathcal{M}(\mathcal{T}_N) = \mathcal{M}(\mathcal{T}_0)$. La valeur commune est l’ensemble vide si et seulement si le tableau est fermé, ou encore si et seulement si φ n’a pas de modèle. Pour être complet, il faut encore s’assurer que le nombre N d’étapes est toujours fini. C’est une conséquence du principe d’induction, puisque toute prolongation ou ramification implique le remplacement d’une formule par des sous-formules.

3.2.6 Ensembles de Hintikka

Un *ensemble de Hintikka* est un ensemble de formules qui vérifie les trois propriétés suivantes :

– Il ne comporte pas de paire complémentaire de littéraux.
– S’il contient une formule de type α, il contient aussi ses composantes α_1 et α_2.

[47] L’étiquette d’un nœud est un ensemble de formule ; une interprétation rend vrai un ensemble de formule si elle rend vraies toutes les formules de l’ensemble.

– S'il contient une formule de type β, il contient aussi au moins une de ses composantes β_1 et β_2.

Il est clair qu'un ensemble de Hintikka est toujours consistant.[48] Toute branche ouverte d'un tableau sémantique est, par construction, un ensemble de Hintikka.

3.3 La méthode analytique des séquents

3.3.1 Introduction

La méthode des tableaux sémantiques admet une méthode duale, dite "des séquents". Nous donnons d'abord l'algorithme qui transforme un tableau sémantique en une dérivation de séquent. La figure 18 présente un tableau sémantique et la dérivation de séquent correspondante.

– On opère une symétrie d'axe horizontal, amenant les feuilles en haut et la racine en bas. (La construction directe d'une dérivation de séquent procède donc de bas en haut.)
– Chaque étiquette du nouvel arbre se compose des négations des éléments de l'étiquette correspondante du tableau sémantique ; de plus, chaque étiquette est précédée du symbole $\rightarrow$.
– Les arcs du tableau deviennent des lignes horizontales dans la dérivation de séquent.
– Les symboles $\bigcirc$ et $\times$ sont remplacés respectivement par les symboles **H** et **A**.

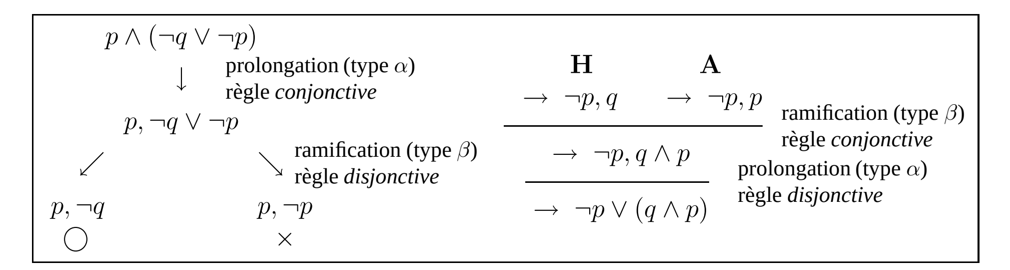


FIG. 18 – Un tableau sémantique et une dérivation de séquent.

Un tableau sémantique et la dérivation de séquent correspondante ne se rapportent donc pas à une même formule. On peut éliminer cette divergence (au moins en apparence) en utilisant la notation signée plutôt que la notation classique pour représenter le tableau (voir figure 19).

3.3.2 Interprétation

Fondamentalement, le contenu sémantique d'une dérivation de séquent est le même que celui du tableau correspondant, mais la dualité permet de présenter ce contenu différemment.

– Chaque étiquette d'une dérivation de séquent s'interprète comme un ensemble *disjonctif* de formules.
– Les feuilles correspondent à des *clauses* c'est-à-dire à des disjonctions de littéraux.

[48] C'est aussi une conséquence du principe d'induction.

$\neg p \vee (q \wedge p) = \mathbf{F}$
prolongation (type α)
règle *conjonctive*
$p = \mathbf{V},\ q \wedge p = \mathbf{F}$
ramification (type β)
règle *disjonctive*
$p = \mathbf{V},\ q = \mathbf{F}$ $\bigcirc$
$p = \mathbf{V},\ p = \mathbf{F}$ $\times$

H $\rightarrow \neg p, q$ **A** $\rightarrow \neg p, p$
ramification (type β) règle *conjonctive*
$\rightarrow \neg p, q \wedge p$
prolongation (type α) règle *disjonctive*
$\rightarrow \neg p \vee (q \wedge p)$

FIG. 19 – Un tableau sémantique signé et une dérivation de séquent.

– Les feuilles valides sont étiquetées **A** ; ce symbole signifie "Axiome" : vérité universelle. Les feuilles non valides sont étiquetées **H** ; ce symbole signifie "Hypothèse" : énoncé contingent.
– La ligne horizontale s'interprète comme la relation d'équivalence logique : une interprétation rend vraie(s) la (les) *prémisse(s)*, au numérateur, si et seulement si elle rend vraie la *conclusion*, au dénominateur.

On a aussi les définitions et règles suivantes.

– Une clause est un *axiome* si elle comporte une paire complémentaire de littéraux, et une *hypothèse* sinon.
– Les *règles d'inférence* sont de deux types
  – *règles* α (prolongation) :

$$\frac{\rightarrow\ U \cup \{\alpha_1, \alpha_2\}}{\rightarrow\ U \cup \{\alpha\}}$$

  – *règles* β (ramification) :

$$\frac{\rightarrow\ U \cup \{\beta_1\} \qquad\qquad \rightarrow\ U \cup \{\beta_2\}}{\rightarrow\ U \cup \{\beta\}}$$

α	α_1	α_2
$A_1 \vee A_2$	A_1	A_2
$\neg(A_1 \wedge A_2)$	$\neg A_1$	$\neg A_2$
$A_1 \Rightarrow A_2$	$\neg A_1$	A_2
$A_1 \Leftarrow A_2$	A_1	$\neg A_2$

β	β_1	β_2
$B_1 \wedge B_2$	B_1	B_2
$\neg(B_1 \vee B_2)$	$\neg B_1$	$\neg B_2$
$\neg(B_1 \Rightarrow B_2)$	B_1	$\neg B_2$
$\neg(B_1 \Leftarrow B_2)$	$\neg B_1$	B_2

FIG. 20 – Les règles de (dé)composition.

Rappelons que les doubles négations sont systématiquement simplifiées et que le biconditionnel et la disjonction exclusive sont interdits. On observe (figure 20) que les formules *conjonctives* donnent lieu à une *ramification* (type β) et les formules *disjonctives* à une *prolongation* (type α). C'était le contraire pour les tableaux sémantiques.

Si on lit la dérivation de haut en bas, les règles de décomposition deviennent des règles de composition, ou *règles d'inférence*.

3.3.3 Propriétés de la méthode des séquents

Elles se déduisent immédiatement de celles des tableaux. La méthode de dérivation de séquent est adéquate : toute formule racine d'une dérivation de séquent dont toutes les feuilles sont des axiomes est valide. La méthode de dérivation de séquent est complète : toute formule valide est racine d'une dérivation de séquent dont toutes les feuilles sont des axiomes.

3.3.4 Extension d'écriture

On convient que le séquent
$\rightarrow \neg A, B, \neg C, \neg D, E, F$
peut aussi s'écrire
$A, C, D \ \rightarrow \ B, E, F.$
Ce séquent est vrai pour v si v rend vraie au moins une des formules B, E et F, ou si v rend fausse au moins une des formules A, C et D. La partie de gauche (ici A, C, D) est l'*antécédent*, la partie de droite (ici B, E, F) est le *succédent*. Le séquent est vrai pour v si et seulement si le conditionnel $(A \wedge C \wedge D) \Rightarrow (B \vee E \vee F)$ est vrai pour v.

La virgule a valeur conjonctive dans l'antécédent et valeur disjonctive dans le succédent. Un antécédent vide correspond à $true$, un succédent vide correspond à $false$, le séquent vide correspond à $false$. Tout séquent dont l'antécédent et le succédent comportent une formule commune est valide.

On peut transférer une formule de l'antécédent au succédent, ou inversement, en changeant sa *polarité*, c'est-à-dire en transformant A en $\neg A$ ou inversement. Le séquent $A, C, D \ \rightarrow \ B, E, F$ est donc logiquement équivalent au séquent $A, \neg B, C \ \rightarrow \ \neg D, E, F$.

Il est commode de récrire les règles en tenant compte des nouvelles notations. A titre d'exemple, si A et B désignent des formules et si U et V désignent des ensembles de formules, les anciennes règles

$$\frac{\rightarrow \ V, \neg A, B}{\rightarrow \ V, (A \Rightarrow B)} \qquad \frac{\rightarrow \ V, A \qquad \rightarrow \ V, \neg B}{\rightarrow \ V, \neg(A \Rightarrow B)}$$

peuvent être étendues en

$$\frac{U \ \rightarrow \ V, \neg A, B}{U \ \rightarrow \ V, (A \Rightarrow B)} \qquad \frac{U \ \rightarrow \ V, A \qquad U \ \rightarrow \ V, \neg B}{U \ \rightarrow \ V, \neg(A \Rightarrow B)}$$

puis récrites en les nouvelles règles suivantes :

$$\frac{U, A \ \rightarrow \ V, B}{U \ \rightarrow \ V, (A \Rightarrow B)} \qquad \frac{U \ \rightarrow \ V, A \qquad U, B \ \rightarrow \ V}{U, (A \Rightarrow B) \ \rightarrow \ V}$$

3.3.5 Règles réversibles, règles analytiques et synthétiques

La règle d'inférence

$$\frac{U \ \rightarrow \ V, A \qquad U, B \ \rightarrow \ V}{U, (A \Rightarrow B) \ \rightarrow \ V}$$

est *réversible* : la barre horizontale peut s'interpréter comme l'*équivalence* logique. Dans une règle non réversible, la conclusion est conséquence logique des prémisses, mais non l'inverse. Si on pose $U_c = \bigwedge U$ et $V_d = \bigvee V$, la règle ci-dessus exprime que les modèles communs des formules $U_c \Rightarrow (V_d \vee A)$ et $(U_c \wedge B) \Rightarrow V_d$ *sont exactement* les modèles de la formule $(U_c \wedge (A \Rightarrow B)) \Rightarrow V_d$.

Cette règle est aussi *analytique* : toute formule apparaissant en haut apparaît aussi en bas (comme formule ou sous-formule). Les dérivations de séquents peuvent se lire de bas en haut (analyse d'une formule) ou de haut en bas (déduction d'une formule au départ d'axiomes et/ou d'hypothèses).

Il existe aussi des méthodes *synthétiques* de déduction, ne se prêtant pas directement à l'analyse des formules. La méthode synthétique est le plus souvent la seule utilisable en mathématique. On utilise des règles où la barre s'interprète comme la relation (non symétrique) de *conséquence* logique. L'exemple le plus connu de règle synthétique (donc non analytique) et non réversible est sans doute le *Modus ponens* :

$$\frac{U \ \rightarrow \ A \qquad U \ \rightarrow \ (A \Rightarrow B)}{U \ \rightarrow \ B}$$

Un exemple de règle synthétique mais réversible est la règle de *coupure* :

$$\frac{U, A \ \rightarrow \ V \qquad U \ \rightarrow \ V, A}{U \ \rightarrow \ V}$$

La formule A et ses sous-formules peuvent ne pas apparaître dans les conclusions. Il est donc difficile de "deviner" des prémisses adéquates au départ des conclusions.[49]

3.3.6 Différences entre conditionnel et séquent

Dans la présentation qui vient d'être faite, la seule différence est que les deux termes d'un séquent sont des ensembles (éventuellement infinis) de formules, tandis que les termes d'un conditionnel sont des formules, finies par nature.

Une autre différence plus subtile apparaît souvent. Les séquents peuvent être utilisés dans des contextes variés, mais le sont habituellement dans un contexte de preuve de validité. Autrement dit, seules des dérivations sans hypothèses sont considérées. Tout séquent apparaissant dans une telle dérivation est valide, et il est alors naturel et fréquent d'introduire cette information dans la définition même de la notion de séquent. Cela signifie qu'un séquent

[49] Les deux règles synthétiques de Modus ponens et de coupure formalisent deux modes de raisonnement omniprésents en mathématique et dans la vie quotidienne. Le Modus ponens traduit la notion même de théorème ou de résultat général : "appliquer" $(A \Rightarrow B)$, c'est déduire B dans le cas où A est connu. La règle de coupure formalise le raisonnement par cas : si on infère V de U quand A est vrai, et aussi quand A est faux, on infère V de U en toute généralité.

n'est plus un objet du langage (assimilable à un conditionnel) mais devient une implication logique, un objet du métalangage. Dans ce cadre, la signification du séquent

$$A, B, C \ \rightarrow \ D, E, F$$

noté parfois

$$A, B, C \vdash D, E, F$$

est "Le conditionnel $(A \wedge B \wedge C) \Rightarrow (D \vee E \vee F)$ est valide".

3.3.7 Tableaux signés vs. séquents

Il y a correspondance naturelle entre les tableaux signés et les séquents : les formules apparaissent dans l'antécédent ou dans le succédent d'un séquent selon qu'elles sont assertées positivement ou négativement dans le nœud homologue du tableau (figure 21).

$$\neg p \vee (q \wedge p) = \mathbf{F}$$
$$\downarrow$$
$$p = \mathbf{V},\ q \wedge p = \mathbf{F}$$
$$\swarrow \qquad \searrow$$
$$p = \mathbf{V},\ q = \mathbf{F} \qquad p = \mathbf{V},\ p = \mathbf{F}$$
$$\bigcirc \qquad \times$$

$$\dfrac{\dfrac{\overset{\mathbf{H}}{p \ \rightarrow \ q} \qquad \overset{\mathbf{A}}{p \ \rightarrow \ p}}{p \ \rightarrow \ q \wedge p}}{\rightarrow \ \neg p \vee (q \wedge p)}$$

FIG. 21 – Un tableau sémantique signé et une dérivation de séquent.

3.4 Le raisonnement automatique

3.4.1 Introduction

La logique formelle permet, comme on vient de le voir tout au long de ce chapitre, de transformer le raisonnement en un objet mathématique, susceptible d'être traité, et même construit, par un ordinateur. Nous avons vu, au paragraphe 2.4.3, qu'un texte court mais relativement dense, présentant un raisonnement, pouvait être transformé en formules et ainsi devenir accessible à une analyse fiable et automatique.[50] Un célèbre ouvrage de science-fiction[51] évoque ainsi l'analyse formelle d'un très compliqué et volumineux document diplomatique … aboutissant à la conclusion que ce document était sémantiquement vide et que ses auteurs méritaient, étymologiquement du moins, leur statut de diplomate. Plus concrètement, un livre d'analyse mathématique a été entièrement vérifié par ordinateur dans le cadre d'un projet d'intelligence artificielle. Peut-on réellement espérer ramener ainsi le raisonnement au calcul, dans le but de l'automatiser ?

[50] L'informaticien dirait : fiable parce que complètement automatique …

[51] *Foundation*, d'Isaac Asimov.

3.4.2 Digression : Leibniz et le raisonnement automatisable

Leibniz avait imaginé un "ratiocinator universalis", sorte de machine abstraite capable de raisonner et, dans une certaine mesure, de trancher des débats épineux. Il croyait possible de représenter les idées et le raisonnement dans un langage symbolique pourvu d'une syntaxe et d'une sémantique précises. Dans une discussion, il devenait théoriquement possible de remplacer un débat d'idées animé par une séance de froid calcul dont l'issue serait inconditionnellement admise par les protagonistes du débat. Boole et surtout Frege ont créé le langage symbolique rêvé par Leibniz, et la logique prédicative est parfaitement apte à la représentation de tout type de raisonnement.[52] Il y a cependant loin entre la capacité de représenter un problème et la capacité de le résoudre ; dans cette section, nous évoquons quelques aspects fragmentaires mais importants de cette question.

3.4.3 Automatiser la logique

Cette question est comme beaucoup d'autres : il est plus facile de l'examiner dans le cadre propositionnel que dans le cadre prédicatif, et certaines conclusions établies dans le cadre propositionnel s'étendront au cadre prédicatif. Cela étant admis, on pourrait croire qu'il n'y a pas de question. La logique propositionnelle est en effet automatisable, puisque nous l'avons effectivement automatisée. La méthode des tables de vérité et celle des tableaux sémantiques par exemple, permettent d'analyser tout raisonnement propositionnel, si complexe soit-il. Nous allons voir maintenant que la portée *pratique* de ces méthodes est sévèrement limitée par un problème de dimension. En dépit de la puissance qu'ils ont atteinte actuellement (et qui continuera à croître de longues années encore), les ordinateurs ne sont pas capables d'analyser, en toute généralité, un problème logique d'une certaine taille. Considérons par exemple le problème classique consistant à déterminer si un ensemble de formules est consistant ou pas. Si nous utilisons les tables de vérité, et si l'ensemble en question comporte des occurrences de n propositions élémentaires distinctes, il suffit de construire une table de vérité … qui comportera 2^n lignes. Si n vaut 30, la table comportera plus d'un milliard de lignes ; si n vaut 100, ce qui n'a rien d'irréaliste, le problème est, sauf cas particulier, définitivement hors d'atteinte de tout ordinateur présent ou à venir. Observons au passage que *multiplier* par 1 000 les performances d'un ordinateur ne permet que d'*ajouter* 10 nouvelles variables propositionnelles à notre lexique.

Il existe a priori deux moyens de contourner cet écueil. D'une part, il est possible de développer des procédures de décision plus rapides que la méthode des tables de vérité et, d'autre part, on peut essayer d'isoler certains types de formules et d'ensembles de formules pour lesquels le problème de la consistance pourrait se résoudre plus rapidement. Nous avons déjà adopté la première approche : la méthode des tableaux sémantiques est souvent — mais pas toujours — nettement plus efficace que celle des tables de vérité. Une analyse plus fine montrerait quand même que cette méthode et, à des degrés divers, toutes les méthodes connues actuellement, restent fondamentalement trop lentes. Plus précisément, dans beaucoup de cas, les performances se dégradent très vite dès que la dimension du problème augmente. Une

[52] Des logiques spéciales ont été et continuent à être introduites pour modéliser plus commodément certains aspects de la connaissance mais, fondamentalement, la logique prédicative peut prétendre à l'universalité.

théorie récemment développée[53] laisse peu de chances de progrès significatifs dans cette voie.

La seconde approche est plus prometteuse ; nous la développons dans la suite de ce chapitre.

3.4.4 Cubes, clauses et formes normales

Considérons la table de vérité de la figure 22, se rapportant à une formule inconnue X, dépendant des trois variables propositionnelles p, q et r. Peut-on reconstituer la formule sur base de la table ? Oui, à une équivalence logique près. La table indique que la formule est vraie dans quatre cas, correspondant aux lignes 1, 3, 5 et 6. A chaque cas correspond un *cube*, c'est-à-dire une conjonction de littéraux. Par exemple, le cube correspondant à la ligne 6 est $(\neg p \wedge q \wedge \neg r)$, ce qui s'interprète en $v(p) = \mathbf{F}$, $v(q) = \mathbf{V}$ et $v(r) = \mathbf{F}$.

p	q	r	$X(p,q,r)$
V	**V**	**V**	**V**
V	**V**	**F**	**F**
V	**F**	**V**	**V**
V	**F**	**F**	**F**
F	**V**	**V**	**V**
F	**V**	**F**	**V**
F	**F**	**V**	**F**
F	**F**	**F**	**F**

FIG. 22 – Table de vérité d'une formule inconnue

La formule X est donc (logiquement équivalente à) la formule

$$(p \wedge q \wedge r) \vee (p \wedge \neg q \wedge r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) .$$

Cette formule peut s'écrire différemment, et notamment

$$(p \wedge r) \vee (\neg p \wedge q) ,$$

ou encore

$$(p \Rightarrow r) \wedge (\neg p \Rightarrow q) .$$

Le point important est que toute formule, puisqu'elle admet une table de vérité, est équivalente à une disjonction de cubes, ce que l'on appelle aussi une *forme disjonctive normale*.

Observons aussi que la négation d'une disjonction de cubes est une conjonction de clauses, ce que l'on appelle aussi une *forme conjonctive normale*. On obtient aisément la forme conjonctive normale d'une formule à partir de la forme disjonctive normale de la négation de cette formule ; on peut aussi l'obtenir directement à partir de la table de vérité, en considérant les lignes pour lesquelles la formule est fausse. Par exemple, la formule $X(p, q, r)$ est fausse dans quatre cas ; elle peut donc s'écrire

$$\neg(p \wedge q \wedge \neg r) \wedge \neg(p \wedge \neg q \wedge \neg r) \wedge \neg(\neg p \wedge \neg q \wedge r) \wedge \neg(\neg p \wedge \neg q \wedge \neg r)$$

[53]Et notamment le théorème de la NP-complétude du problème de la consistance, démontré par Cook en 1970.

ou encore

$$(\neg p \vee \neg q \vee r) \wedge (\neg p \vee q \vee r) \wedge (p \vee q \vee \neg r) \wedge (p \vee q \vee r) ,$$

ce qui est une forme conjonctive normale, c'est-à-dire une conjonction de clauses ; la formule peut se simplifier en

$$(\neg p \vee r) \wedge (p \vee q) ,$$

qui est encore une forme conjonctive normale, appelée aussi *forme clausale*.

Le principe de la déduction affirme que la formule A est conséquence logique de l'ensemble E si et seulement si l'ensemble $E \cup \{\neg A\}$ est inconsistant. Chaque formule de ce dernier ensemble est logiquement équivalente à une conjonction de clauses ; si $\mathcal{C}$ est l'ensemble de toutes ces clauses, on peut dire que A est conséquence logique de E si et seulement si $\mathcal{C}$ est inconsistant. Le problème fondamental de la logique se résume donc à celui de déterminer si un ensemble de clauses est inconsistant ou non.

Remarque. Construire la table de vérité d'une formule donnée n'est généralement pas le moyen le plus rapide d'obtenir une forme normale disjonctive ou conjonctive logiquement équivalente à cette formule.

3.4.5 Clauses de Horn et ensembles de Horn

Nous venons de rappeler que, pour les problèmes d'une certaine taille, l'approche automatique était très aléatoire. C'est étonnant, dans la mesure où l'être humain moyen est capable d'effectuer des raisonnements logiques de grande taille avec une certaine efficacité. Une raison à cela est l'aptitude, typique de l'être humain, à s'adapter aux circonstances et à remplacer une méthode générale par une approche plus spécifique et plus rapide, du moins dans le cas particulier considéré. Une autre raison, plus pertinente ici, est que bien souvent les formules de l'ensemble E, qui constituent la “base de connaissance” à partir de laquelle on va déduire la formule A, ont une forme très particulière, qui se retrouve aussi dans les énoncés mathématiques, à savoir

$$(p_1 \wedge \cdots \wedge p_n) \Rightarrow q .$$

Les p_i et q sont des propositions élémentaires.[54] La formule exprime simplement que toute interprétation rendant vraies les propositions $p_1, \ldots, p_n$ rend vraie aussi la conclusion q.[55]

On voit immédiatement que cette formule est une clause, que l'on peut récrire en

$$\neg p_1 \vee \cdots \vee \neg p_n \vee q .$$

De même, la conclusion A est souvent une simple proposition, ou une conjonction de propositions, donc de litéraux positifs. Sa négation est une disjonction de littéraux négatifs.

On appelle *clause de Horn* toute clause contenant au plus un littéral positif. Une clause de Horn est *définie* si elle comporte exactement un littéral positif ; elle est *négative* si elle n'en comporte pas. Un *ensemble de Horn* est un ensemble de clauses de Horn. Cette notion est extrêmement importante parce que, dans le cas particulier des ensembles de Horn, le problème de la consistance peut se résoudre de manière efficace, en toute généralité.

[54]On peut considérer que cette formule représente un théorème mathématique, dont les p_i sont les hypothèses et q la thèse.

[55]On observera que cette formule, même si elle représente un théorème de mathématique, n'est pas valide. La raison en est que dans le cadre de la logique propositionnelle, les propositions élémentaires ne sont pas analysées.

3.4.6 L'algorithme de résolution unitaire

Une *clause unitaire* est une clause comportant un seul littéral. Une clause unitaire peut être *positive* ou *négative*. Dans la suite, sauf mention explicite du contraire, une clause unitaire est une clause unitaire positive, réduite à une proposition élémentaire. L'algorithme de *résolution unitaire* (figure 23) est un moyen simple de déterminer si un ensemble de Horn est consistant ou inconsistant.

$\{S := S_0\}$
Tant que $\Box \notin S$ faire
  choisir p et c tels que
    p est une clause unitaire positive de S,
    c est une clause de S contenant $\neg p$;
  $r := c \setminus \{\neg p\}$;
  $S := (S \setminus \{c\}) \cup \{r\}$.

FIG. 23 – Résolution unitaire

Le principe de cet algorithme est très simple. Il consiste à supprimer, dans les clauses d'un ensemble S de clauses de Horn, tous les littéraux négatifs dont la proposition sous-jacente apparaît comme clause unitaire, jusqu'à ce qu'une clause soit devenue vide, ou que plus aucune suppression ne soit possible. Dans le premier cas, on conclut à l'inconsistance, dans le second, à la consistance. La notation "$:=$" signifie "devient" ; exécuter l'instruction $x := x + y$ signifie que la nouvelle valeur de x est l'ancienne valeur de $x + y$. Si c est une clause contenant $\neg p$, $c \setminus \{\neg p\}$ est la clause obtenue en supprimant $\neg p$. La *clause vide*, qui ne contient aucun littéral, est représentée par le symbole $\Box$. Une clause est vraie si au moins un de ses littéraux est vrai ; la clause vide est donc logiquement équivalente à false. Si c est une clause de l'ensemble S, $S \setminus \{c\}$ est l'ensemble obtenu en enlevant c de S. Si r est une clause, $S \cup \{r\}$ est l'ensemble obtenu en ajoutant r à S.

La figure 24 donne un exemple d'exécution de l'algorithme, permettant de montrer que l'ensemble

$$S = \{p \vee \neg r \vee \neg t\, ,\ q\, ,\ r\, ,\ t \vee \neg p \vee \neg r\, ,\ t \vee \neg q\, ,\ \neg p \vee \neg q \vee \neg r\}$$

est inconsistant.

1.	$p \vee \neg r \vee \neg t$	$\underline{q}$	r	$t \vee \neg p \vee \neg r$	$t \vee \underline{\neg q}$	$\neg p \vee \neg q \vee \neg r$
2.	$p \vee \underline{\neg r} \vee \neg t$	q	$\underline{r}$	$t \vee \neg p \vee \neg r$	t	$\neg p \vee \neg q \vee \neg r$
3.	$p \vee \neg t$	$\underline{q}$	r	$t \vee \neg p \vee \neg r$	t	$\neg p \vee \underline{\neg q} \vee \neg r$
4.	$p \vee \neg t$	q	$\underline{r}$	$t \vee \neg p \vee \neg r$	t	$\neg p \vee \underline{\neg r}$
5.	$p \vee \underline{\neg t}$	q	r	$t \vee \neg p \vee \neg r$	$\underline{t}$	$\neg p$
6.	$\underline{p}$	q	r	$t \vee \neg p \vee \neg r$	t	$\underline{\neg p}$
7.	p	q	r	$t \vee \neg p \vee \neg r$	t	$\Box$

FIG. 24 – Résolution unitaire : un exemple positif

La figure 25 donne un second exemple d'exécution de l'algorithme, permettant de montrer que l'ensemble

$$S = \{p \vee \neg r \vee \neg t\, ,\ q\, ,\ s\, ,\ t \vee \neg p \vee \neg r\, ,\ t \vee \neg q \vee \neg s\, ,\ \neg p \vee \neg q \vee \neg r\}$$

est consistant.

1.	$p \vee \neg r \vee \neg t$	$\underline{q}$	s	$t \vee \neg p \vee \neg r$	$t \vee \underline{\neg q} \vee \neg s$	$\neg p \vee \neg q \vee \neg r$
2.	$p \vee \neg r \vee \neg t$	q	$\underline{s}$	$t \vee \neg p \vee \neg r$	$t \vee \underline{\neg s}$	$\neg p \vee \neg q \vee \neg r$
3.	$p \vee \neg r \vee \neg t$	$\underline{q}$	s	$t \vee \neg p \vee \neg r$	t	$\neg p \vee \underline{\neg q} \vee \neg r$
4.	$p \vee \neg r \vee \underline{\neg t}$	q	s	$t \vee \neg p \vee \neg r$	$\underline{t}$	$\neg p \vee \neg r$
5.	$p \vee \neg r$	q	s	$t \vee \neg p \vee \neg r$	t	$\neg p \vee \neg r$

FIG. 25 – Résolution unitaire : un exemple négatif

La première propriété de l'algorithme de résolution unitaire est d'être efficace. A chaque étape, un littéral est enlevé donc le nombre d'étapes ne peut dépasser le nombre total de littéraux. A chaque étape, l'ensemble S change (un littéral est supprimé). Comme d'habitude, le point crucial de l'analyse de l'algorithme consiste à déterminer ce qui ne change pas. Comme précédemment, c'est l'ensemble des modèles de S qui ne change pas. En effet, si la clause unitaire p se trouve dans S, seules les interprétations rendant p vrai peuvent être des modèles. Soit I une telle interprétation ; quelle que soit la clause c contenant $\neg p$, on a $I(c) = I(c \setminus \{\neg p\})$. On a donc, après chaque étape, $\mathcal{M}(S) = \mathcal{M}(S_0)$. En particulier, après la dernière étape, on a $\mathcal{M}(S_f) = \mathcal{M}(S_0)$, si S_f désigne l'état final de l'ensemble S. Cela prouve en particulier que S_0 (l'ensemble de départ, celui qui nous intéresse) est consistant si et seulement si S_f est consistant. Il se fait que déterminer si S_f est consistant est immédiat. En effet, il n'y a que deux possibilités :

– L'ensemble S_f contient la clause vide, qui est inconsistante, donc S_f est inconsistant.
– L'ensemble S_f ne contient pas la clause vide. Soit I l'interprétation qui rend vraies toutes les clauses unitaires (positives) de S_f et fausses toutes les autres propositions. Cette interprétation rend vraies toutes les clauses unitaires, et aussi toutes les clauses non unitaires, puisque ces dernières contiennent au moins un littéral négatif dont la proposition sous-jacente est fausse par définition de I (car cette proposition n'est pas une clause unitaire, sinon elle donnerait lieu à une étape supplémentaire).

On appelle *base de connaissance* un ensemble de clauses de Horn positives ; on appelle *question* une conjonction de propositions. L'algorithme de résolution unitaire permet de déterminer si une question A est conséquence logique d'une base de connaissance H ; ce sera le cas si l'ensemble $H \cup \{\neg A\}$ est reconnu inconsistant.

Remarque. On peut aussi tester l'ensemble H seul ; il est nécessairement consistant puisqu'il ne comporte que des clauses de Horn positives.[56] La détermination de H_f permet d'obtenir l'ensemble de toutes les propositions qui sont conséquences logiques de H ; ce sont les propositions qui apparaissent comme clauses unitaires dans H_f. L'interprétation qui rend vraies ces propositions et seulement celles-là est le *modèle canonique*, ou *modèle minimal* de H. Le

[56]L'interprétation qui rend vraies toutes les propositions est donc un modèle de H.

modèle minimal de l'ensemble traité à la figure 25 est donc l'interprétation qui rend vraies les propositions q, s et t et seulement celles-là.

Remarque. On représente souvent les exécutions de l'algorithme de résolution unitaire sous forme arborescente ; la représentation correspondant à l'exécution de la figure 24 se trouve à la figure 26. L'arborescence s'appelle *arbre de dérivation*, ou *arbre de réfutation* dans le cas particulier où on dérive la clause vide.

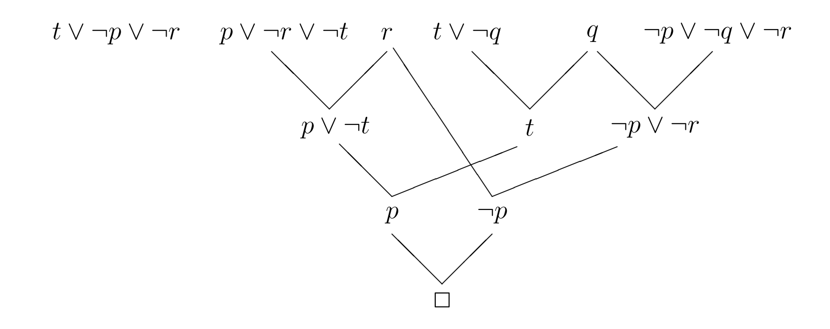


FIG. 26 – Arbre de réfutation unitaire pour l'ensemble S

3.4.7 La programmation logique propositionnelle

Le problème de la programmation logique propositionnelle consiste à déterminer si une proposition est ou n'est pas conséquence logique d'un ensemble de clauses de Horn définies, appelé *base de connaissance* ou *programme logique*. Nous venons de voir que l'algorithme de résolution unitaire est une solution générale et raisonnablement efficace pour ce problème. Elle n'est cependant pas optimale en pratique. La raison en est que, dans la plupart des cas, la base de connaissance est énorme, voire infinie, et que la plupart des clauses qu'elle contient n'ont rien à voir avec la question particulière à traiter. L'algorithme de résolution unitaire n'accorde aucun rôle particulier à la question traitée, dont la négation est simplement ajoutée à la base de connaissance. L'*algorithme de résolution d'entrée* est une variante de l'algorithme de résolution unitaire, dans laquelle la négation de la question joue un rôle privilégié. Cette variante est représentée à la figure 27, où L désigne un programme logique.

$\{G = G_0\}$
Tant que $G \neq \Box$ faire
  choisir p et c tels que
    $\neg p \in G$,
    $c \in L$ et
    $p \in c$;
  $G := (G \setminus \{\neg p\}) \vee (c \setminus \{p\})$.

FIG. 27 – Résolution d'entrée

Cet algorithme utilise une variable G, appelée le *but*, dont la valeur est toujours une clause de Horn négative ; initialement, le but est la négation de la question. A chaque étape, le but est transformé selon la règle suivante : un littéral $\neg p$ du but est remplacé par $(c \setminus \{p\})$, où c est une clause de la base de connaissance, dont le littéral positif est p. On pourrait démontrer que l'algorithme de résolution d'entrée est équivalent à l'algorithme de résolution unitaire. A titre d'exemple, nous montrons à la figure 28 que la proposition p est bien conséquence logique du programme logique

$$L = \{\, t \vee \neg p \vee \neg r \,,\ p \vee \neg r \vee \neg t \,,\ r \,,\ t \vee \neg q \,,\ q \,\} \,.$$

FIG. 28 – Arbre de réfutation d'entrée pour L et p

On voit que, dans un arbre de réfutation d'entrée, il existe une branche principale unissant le but initial (ici, $\neg p$) à la clause vide. Les branches auxiliaires sont de longueur 1 et unissent une clause d'entrée (d'où le nom de l'algorithme) à un but intermédiaire.

3.4.8 Prolog propositionnel

L'algorithme de Prolog est une version concrète de l'algorithme de résolution d'entrée. Les clauses sont représentées par des listes de littéraux et le programme logique L est une liste de clauses. Les choix de p et c sont imposés par une stratégie très simple ; p est nécessairement la proposition correspondant au premier littéral du but et c est la première clause de L convenable. En effet, le système Prolog essaiera, dans l'ordre où elles se présentent dans la liste L, toutes les clauses dont la tête est p. Cela peut poser un problème si l'ordre des clauses ou des littéraux dans une clause est mal choisi. A titre d'exemple, voici la version Prolog du programme logique L donné plus haut :

```
t :- p,r.
p :- r,t.
r.
t :- q.
q.
```

On voit que la virgule a valeur conjonctive et que le symbole "`:-`" représente le connecteur $\Leftarrow$ (conditionnel inverse). L'ordre n'est pas adéquat ici car la clause inutile "`t :- p,r.`" court-circuite la clause utile "`t :- q.`". Si on omet la clause inutile, le système Prolog détecte que la proposition p est conséquence logique du programme logique L. Nous reviendrons sur le système Prolog dans le cadre prédicatif, qui permet des applications plus intéressantes.

3.5 Quelques exercices

3.5.1 Argumentation

Le récit de la création du monde. Au paragraphe 2.4.3, nous avons formalisé un argument ; les techniques présentées dans ce chapitre permettent de déterminer si cet argument est correct ou non ou, plus précisément, si l'argument formel correspondant est correct. Un argument formel se compose d'un ensemble (fini) de prémisses $E = \{P_1, \ldots, P_n\}$ et d'une conclusion C ; il est dit *correct* si la conclusion est conséquence logique des prémisses, ce qui s'écrit $E \models C$; cela a lieu si et seulement si le conditionnel $(P_1 \wedge \cdots \wedge P_n) \Rightarrow C$ est valide. L'argument formel introduit au paragraphe 2.4.3 conduit donc à la question

$$\{E \Rightarrow \neg Q\,,\ \neg Q \Rightarrow \neg A\,,\ D \vee A\} \overset{?}{\models} \neg E \vee D\,.$$

La méthode des tables de vérité est peu intéressante ici car le lexique utilisé comporte quatre propositions ; la table de vérité aurait donc seize lignes. La méthode des tableaux sémantiques peut s'appliquer ; si nous croyons que l'argument est correct, la racine de notre tableau sera la négation du conditionnel associé à cet argument. La figure 29 représente ce tableau.

$$\neg([(E \Rightarrow \neg Q) \wedge (\neg Q \Rightarrow \neg A) \wedge (D \vee A)] \Rightarrow (\neg E \vee D))$$
$$[(E \Rightarrow \neg Q) \wedge (\neg Q \Rightarrow \neg A) \wedge (D \vee A)]\,,\ \neg(\neg E \vee D)$$
$$(E \Rightarrow \neg Q)\,,\ (\neg Q \Rightarrow \neg A)\,,\ (D \vee A)\,,\ \neg(\neg E \vee D)$$
$$\underline{(E \Rightarrow \neg Q)}\,,\ (\neg Q \Rightarrow \neg A)\,,\ (D \vee A)\,,\ E\,,\ \neg D$$
$$\neg E, \ldots, E \ldots \quad \times \qquad\qquad \neg Q\,, \underline{(\neg Q \Rightarrow \neg A)}\,, (D \vee A)\,, E\,, \neg D$$
$$\neg Q\,, Q\,, \ldots \quad \times \qquad\qquad \neg Q\,, \neg A\,, \underline{(D \vee A)}\,, E\,, \neg D$$
$$\ldots, D, \ldots, \neg D \quad \times \qquad\qquad \ldots, \neg A, A, \ldots \quad \times$$

FIG. 29 – Tableau sémantique, argument "biblique"

Ce tableau est fermé, donc l'argument est correct. On notera cependant que cette construction n'est que marginalement moins fastidieuse que celle d'une table de vérité à seize lignes. Il serait souhaitable de disposer de techniques plus expéditives, non seulement pour gagner du temps, mais aussi pour limiter le risque d'erreur.[57] En relisant attentivement la justification de la méthode des tableaux sémantiques, on peut observer que l'étiquette d'un nœud peut être remplacée par une autre, pour peu que les deux étiquettes admettent exactement les mêmes modèles.[58]

La figure 30 donne un tableau sémantique exploitant ce principe. Les simplifications successives se basent sur les faits suivants :

- Les ensembles $\{E \Rightarrow \neg Q\,,\ E\}$ et $\{\neg Q\,,\ E\}$ sont logiquement équivalents ;
- Les ensembles $\{D \vee A\,,\ \neg D\}$ et $\{A\,,\ \neg D\}$ sont logiquement équivalents ;
- Les ensembles $\{\neg Q \Rightarrow \neg A\,,\ A\}$ et $\{Q\,,\ A\}$ sont logiquement équivalents.

$$\neg([(E \Rightarrow \neg Q) \wedge (\neg Q \Rightarrow \neg A) \wedge (D \vee A)] \Rightarrow (\neg E \vee D))$$
$$\downarrow$$
$$[(E \Rightarrow \neg Q) \wedge (\neg Q \Rightarrow \neg A) \wedge (D \vee A)]\,,\ \neg(\neg E \vee D)$$
$$\downarrow$$
$$(E \Rightarrow \neg Q)\,,\ (\neg Q \Rightarrow \neg A)\,,\ (D \vee A)\,,\ \neg(\neg E \vee D)$$
$$\downarrow$$
$$\underline{(E \Rightarrow \neg Q)}\,,\ (\neg Q \Rightarrow \neg A)\,,\ (D \vee A)\,,\ \underline{E}\,,\ \neg D$$
$$\downarrow$$
$$\neg Q\,,\ (\neg Q \Rightarrow \neg A)\,,\ \underline{(D \vee A)}\,,\ E\,,\ \underline{\neg D}$$
$$\downarrow$$
$$\neg Q\,,\ \underline{(\neg Q \Rightarrow \neg A)}\,,\ \underline{A}\,,\ E\,,\ \neg D$$
$$\downarrow$$
$$\neg Q\,,\ Q\,,\ A\,,\ E\,,\ \neg D$$
$$\times$$

FIG. 30 – Tableau sémantique simplifié

Chacune de ces simplifications a permis l'économie d'un branchement.[59]

On peut aussi envisager l'utilisation de la théorie de Horn, pour tester l'inconsistance de l'ensemble composé des prémisses et de la négation de la conclusion :

$$\{E \Rightarrow \neg Q\,,\ \neg Q \Rightarrow \neg A\,,\ D \vee A\,,\ \neg(\neg E \vee D)\}\,,$$

qui se récrit en

$$\{\,\neg E \vee \neg Q\,,\ Q \vee \neg A\,,\ D \vee A\,,\ E\,,\ \neg D\,\}\,.$$

[57] On dit parfois que le taux d'erreur d'un développement formel (non vérifié par ordinateur) est proportionnel au *carré* de la taille de ce développement...

[58] De plus, pour éviter le risque de non-terminaison, la nouvelle étiquette ne pourra être plus complexe que l'ancienne.

[59] Une autre manière de justifier ces simplifications est d'observer que chaque couple simplifiable comporte une formule disjonctive et un littéral, et que la décomposition de la formule disjonctive donne lieu à une branche comportant le littéral opposé et à une branche comportant le couple simplifié. Le "raccourci" proposé consiste à faire l'économie de la première branche, inutile puisque fermable immédiatement.

L'une des clauses de cet ensemble comporte deux littéraux positifs, ce qui est incompatible avec l'utilisation de la théorie de Horn. Dans le cas présent, on remédie à ce problème par obversion.[60] On introduit donc la proposition $\tilde{D}$, définie comme logiquement équivalente à la formule $\neg D$. L'ensemble devient

$$\{\neg E \vee \neg Q\,,\ Q \vee \neg A\,,\ \neg\tilde{D} \vee A\,,\ E\,,\ \tilde{D}\}\,.$$

Cet ensemble de Horn est bien inconsistant, comme le montre le développement de la figure 31.

1.	$\neg E \vee \neg Q$	$Q \vee \neg A$	$\neg\tilde{D} \vee A$	$\underline{E}$	$\tilde{D}$
2.	$\neg Q$	$Q \vee \neg A$	$\neg\tilde{D} \vee A$	E	$\underline{\tilde{D}}$
3.	$\neg Q$	$Q \vee \neg A$	$\underline{A}$	E	$\tilde{D}$
3.	$\neg Q$	$\underline{Q}$	A	E	$\tilde{D}$
4.	$\square$	Q	A	E	$\tilde{D}$

FIG. 31 – Résolution unitaire : argument "biblique"

Croire aux fantômes ? C'est ce que nous suggère le raisonnement ci-dessous, qu'il est prudent d'analyser . . .

> *Si on considère que les gens qui étudient les perceptions extra-sensorielles sont honnêtes, alors il faut admettre l'existence de telles perceptions. De plus, si l'on met à l'épreuve l'existence des perceptions extra-sensorielles, on se doit de considérer sérieusement la doctrine de la clairvoyance. Admettre l'existence des perceptions extra-sensorielles doit nous pousser à mettre celles-ci à l'épreuve et à les expliquer.*
>
> *La doctrine de la clairvoyance doit être considérée sérieusement si on est prêt à considérer sérieusement les phénomènes occultes. Et si on est prêt à considérer sérieusement ces phénomènes, nous devrons respecter les médiums. Aussi, si nous respectons ces gens, nous devrons aussi prendre au sérieux leur prétendue aptitude à communiquer avec les morts. Enfin, si nous devons prendre au sérieux cette aptitude à communiquer avec les morts, on ne peut que croire aux fantômes.*
>
> *Considérer que les gens qui étudient les perceptions extra-sensorielles sont honnêtes nous oblige donc à croire aux fantômes.*

On utilise le lexique suivant :

[60]L'*obversion* consiste à introduire ou à supprimer une négation dans une phrase sans en changer le sens. Par exemple, la phrase "Tout ensemble contenant une paire complémentaire de littéraux est inconsistant" devient par obversion "Aucun ensemble contenant une paire complémentaire de littéraux n'est consistant".

- *hon* on considère que les gens qui étudient les perceptions extra-sensorielles sont honnêtes ;
- *adm* on admet l'existence des perceptions extra-sensorielles ;
- *epr* on met à l'*épr*euve l'existence des perceptions extra-sensorielles ;
- *cla* on considère sérieusement la doctrine de la clairvoyance ;
- *exp* on cherche à expliquer les perceptions extra-sensorielles ;
- *occ* on considère sérieusement les phénomènes occultes ;
- *med* on respecte les *méd*iums ;
- *com* on prend au sérieux l'aptitude des médiums à communiquer avec les morts ;
- *fan* on croit aux fantômes.

Les prémisses sont, pour le premier paragraphe,

$$hon \Rightarrow adm\,,\ epr \Rightarrow cla\,,\ adm \Rightarrow (epr \wedge exp)\,;$$

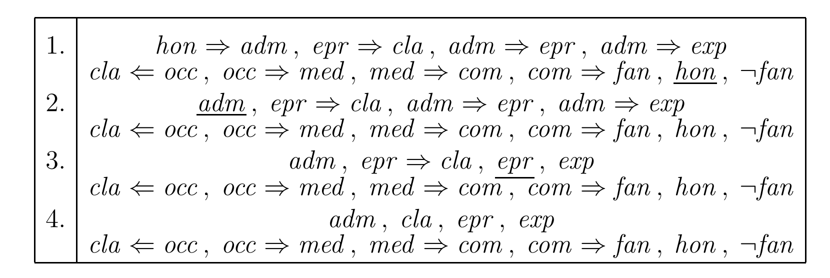
celles du second paragraphe sont

$$cla \Leftarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,.$$

La conclusion (dernier paragraphe) est

$$hon \Rightarrow fan\,.$$

Les procédés utilisés pour résoudre le problème du récit biblique s'appliquent également à ce problème ; nous considérons ici seulement la résolution unitaire. La prémisse $adm \Rightarrow (epr \wedge exp)$ n'est pas une clause, mais est logiquement équivalente à la conjonction des deux clauses $adm \Rightarrow epr$ et $adm \Rightarrow exp$; de même, la négation de la conclusion $hon \Rightarrow fan$ n'est pas une clause, mais est logiquement équivalente à la conjonction des deux clauses hon et $\neg fan$. On obtient ainsi le développement de la figure 32, dans laquelle les clauses (de Horn) ont gardé leur forme conditionnelle.



1.	$hon \Rightarrow adm\,,\ epr \Rightarrow cla\,,\ adm \Rightarrow epr\,,\ adm \Rightarrow exp$ $cla \Leftarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ \underline{hon}\,,\ \neg fan$
2.	$\underline{adm}\,,\ epr \Rightarrow cla\,,\ adm \Rightarrow epr\,,\ adm \Rightarrow exp$ $cla \Leftarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
3.	$adm\,,\ epr \Rightarrow cla\,,\ \underline{epr}\,,\ exp$ $cla \Leftarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
4.	$adm\,,\ cla\,,\ epr\,,\ exp$ $cla \Leftarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$

FIG. 32 – Résolution unitaire : croire au fantômes ?

On voit immédiatement que l'obtention de la nouvelle clause unitaire cla ne permet pas de progresser, car l'unique autre occurrence de cla est positive, dans la prémisse $cla \Leftarrow occ$. Cependant, si la prémisse

> *La doctrine de la clairvoyance doit être considérée sérieusement si on est prêt à considérer sérieusement les phénomènes occultes.*

était remplacée par la prémisse

> *La doctrine de la clairvoyance doit être considérée sérieusement* **seulement** *si on est prêt à considérer sérieusement les phénomènes occultes.*

ou encore, ce qui revient au même, par la prémisse

> *Si la doctrine de la clairvoyance est considérée sérieusement, alors on doit aussi considérer sérieusement les phénomènes occultes.*

la clause $cla \Leftarrow occ$ serait remplacée par la clause $cla \Rightarrow occ$; cela rendrait l'argument correct, comme le montre le développement de la figure 33. On voit toute l'importance qu'un seul mot peut avoir dans un texte . . .

1.	$hon \Rightarrow adm\,,\ epr \Rightarrow cla\,,\ adm \Rightarrow epr\,,\ adm \Rightarrow exp$ $cla \Rightarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ \underline{hon}\,,\ \neg fan$
2.	$\underline{adm}\,,\ epr \Rightarrow cla\,,\ adm \Rightarrow epr\,,\ adm \Rightarrow exp$ $cla \Rightarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
3.	$adm\,,\ epr \Rightarrow cla\,,\ \underline{epr}\,,\ exp$ $cla \Rightarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
4.	$adm\,,\ \underline{cla}\,,\ epr\,,\ exp$ $cla \Rightarrow occ\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
5.	$adm\,,\ cla\,,\ epr\,,\ exp$ $\underline{occ}\,,\ occ \Rightarrow med\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
6.	$adm\,,\ cla\,,\ epr\,,\ exp$ $occ\,,\ \underline{med}\,,\ med \Rightarrow com\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
7.	$adm\,,\ cla\,,\ epr\,,\ exp$ $occ\,,\ med\,,\ \underline{com}\,,\ com \Rightarrow fan\,,\ hon\,,\ \neg fan$
8.	$adm\,,\ cla\,,\ epr\,,\ exp\,,\ occ\,,\ med\,,\ com\,,\ \underline{fan}\,,\ hon\,,\ \neg fan$
9.	$adm\,,\ cla\,,\ epr\,,\ exp\,,\ occ\,,\ med\,,\ com\,,\ fan\,,\ hon\,,\ \Box$

FIG. 33 – Argument "Croire au fantômes ?" corrigé

3.5.2 Analyse de formules

Soient A, B, X, Y des formules. On suppose $A \models B$. Dans les quatre cas suivants, peut-on affirmer $C_i \models D_i$ et/ou $D_i \models C_i$?

1. $C_1 =_{def} X \Rightarrow (A \Rightarrow Y)$ et $D_1 =_{def} X \Rightarrow (B \Rightarrow Y)$;
2. $C_2 =_{def} (X \Rightarrow A) \vee Y$ et $D_2 =_{def} (X \Rightarrow B) \vee Y$;
3. $C_3 =_{def} (X \equiv A) \Rightarrow Y$ et $D_3 =_{def} (X \equiv B) \Rightarrow Y$;
4. $C_4 =_{def} X \equiv ((A \Rightarrow (B \vee A)) \Rightarrow Y)$ et $D_4 =_{def} X \equiv ((B \Rightarrow (A \vee B)) \Rightarrow Y)$.

Comme toujours, la méthode des tables de vérité peut être utilisée. En principe, puisque les formules C et D dépendent des quatre formules A, B, X, Y, seize lignes sont nécessaires. Cependant, l'hypothèse $A \models B$ élimine les cas $A = \mathbf{V}$, $B = \mathbf{F}$, ce qui ne laisse subsister que douze lignes. La table complète est représentée à la figure 34 ; on en déduit immédiatement les solutions :

1. $C_1 \not\models D_1$ et $D_1 \models C_1$;
2. $C_2 \models D_2$ et $D_2 \not\models C_2$;
3. $C_3 \not\models D_3$ et $D_3 \not\models C_3$;
4. $C_4 \models D_4$ et $D_4 \models C_4$.

A	B	X	Y	C_1	D_1	C_2	D_2	C_3	D_3	C_4	D_4
V	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**
V	**V**	**V**	**F**	**F**	**F**	**V**	**V**	**F**	**F**	**F**	**F**
V	**V**	**F**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**
V	**V**	**F**	**F**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**
F	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**
F	**V**	**V**	**F**	**V**	**F**	**F**	**V**	**V**	**F**	**F**	**F**
F	**V**	**F**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**
F	**V**	**F**	**F**	**V**	**V**	**V**	**V**	**F**	**V**	**V**	**V**
F	**F**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**V**
F	**F**	**V**	**F**	**V**	**V**	**F**	**F**	**V**	**V**	**F**	**F**
F	**F**	**F**	**V**	**V**	**V**	**V**	**V**	**V**	**V**	**F**	**F**
F	**F**	**F**	**F**	**V**	**V**	**V**	**V**	**F**	**F**	**V**	**V**

FIG. 34 – Analyse de formules par table de vérité

Remarque. Il se peut que, pour des choix particuliers de A, B, X, Y, on ait $C_i \models D_i$ et/ou $D_i \models C_i$ même si c'est faux dans le cas général. Plus concrètement, le résultat négatif $C_1 \not\models D_1$ que nous venons d'obtenir signifie que pour certains choix des formules A, B, C, D, la condition $A \models B$ ne garantit pas $C_1 \models D_1$; c'est le cas notamment si A et Y sont identiquement vraies et si B et X sont identiquement fausses. Cela n'exclut pas que, pour d'autres choix (par exemple celui où les quatre formules sont identiquement vraies), $C_1 \models D_1$ puisse avoir lieu. En revanche, tout résultat positif, tel $D_1 \models C_1$ ou $C_4 \models D_4$, s'entend pour tous les choix de formules compatibles avec l'hypothèse. Dans le même ordre d'idée, rappelons que les énoncés $\not\models (A \Rightarrow B)$ et $\models \neg(A \Rightarrow B)$ ne sont pas équivalents : le second (qui garantit la validité de A et l'inconsistance de B) est nettement plus fort que le premier.

La méthode des tableaux sémantiques est également utilisable ici. A titre d'exemple, le tableau de la figure 35 montre la consistance de la formule $(A \Rightarrow B) \wedge \neg(D_2 \Rightarrow C_2)$, ce qui établit le résultat négatif $D_2 \not\models C_2$.

Il convient de souligner que les méthodes des tables de vérité et des tableaux sémantiques sont toujours utilisables, mais rarement optimales. Dans le cas présent, on peut arriver aux conclusions plus rapidement, en notant qu'a priori il est évident que certaines interprétations rendent C_i et D_i logiquement équivalentes. De telles interprétations ne doivent naturellement pas être étudiées, puisque nous cherchons à mettre en évidence les differences sémantiques entre les deux formules. Par exemple, C_1 et D_1 sont toujours vraies (et donc logiquement équivalentes) dès que X est fausse ou que Y est vraie ; le problème relatif à C_1 et D_1 peut donc se réduire au problème relatif à $C_1' =_{def} \neg A$ et $D_1' =_{def} \neg B$, puisque C_1 se réduit à C_1', et

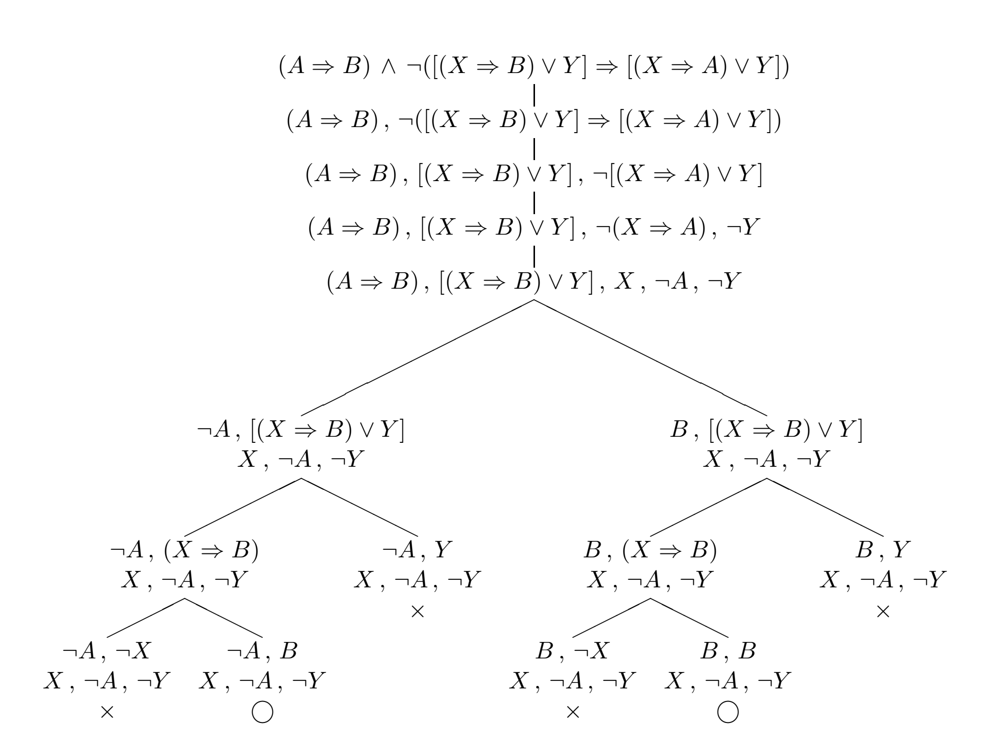


FIG. 35 – Tableau sémantique : $D_2 \not\models C_2$

D_1 se réduit à D'_1 dès que X est vraie et que Y est fausse. Dans le même ordre d'idée, on note que les formules $A \Rightarrow (B \vee A)$ et $B \Rightarrow (A \vee B)$ sont identiquement vraies (valides). En fait, le problème initial se réduit de la sorte au problème concernant les paires suivantes :

1. $C'_1 =_{def} \neg A$ et $D'_1 =_{def} \neg B$;
2. $C'_2 =_{def} A$ et $D'_2 =_{def} B$;
3. $C'_3 =_{def} \neg(X \equiv A)$ et $D'_3 =_{def} \neg(X \equiv B)$;
4. $C'_4 =_{def} X \equiv Y$ et $D'_4 =_{def} X \equiv Y$.

Cette simplification du problème rend les résultats évidents.

Enfin, notons que le tableau sémantique de la figure 35 pouvait être simplifié, comme dans le cas de l'argument "biblique" ; le résultat est donné à la figure 36.

On notera l'emploi d'une nouvelle règle de simplification : la paire $\{(A \Rightarrow B)\,,\, \neg A\}$ est récrite en le singleton $\{\neg A\}$, ces deux ensembles étant logiquement équivalents.

3.5.3 Problèmes

Le coffre partagé. Cinq personnes (A, B, C, D, E) ont des économies en commun dans un coffre. N'ayant pas confiance l'une en l'autre, elles décident que le coffre ne pourra s'ouvrir qu'en présence de A et B, ou de A et C, ou de B, D et E. Combien de serrures le coffre doit-il avoir ? Combien faut-il de clés et à qui les donne-t-on ?

$(A \Rightarrow B) \wedge \neg([(X \Rightarrow B) \vee Y] \Rightarrow [(X \Rightarrow A) \vee Y])$

$(A \Rightarrow B)\,,\ \neg([(X \Rightarrow B) \vee Y] \Rightarrow [(X \Rightarrow A) \vee Y])$

$(A \Rightarrow B)\,,\ [(X \Rightarrow B) \vee Y]\,,\ \neg[(X \Rightarrow A) \vee Y]$

$(A \Rightarrow B)\,,\ [(X \Rightarrow B) \vee Y]\,,\ \neg(X \Rightarrow A)\,,\ \neg Y$

$(A \Rightarrow B)\,,\ (X \Rightarrow B)\,,\ \neg(X \Rightarrow A)\,,\ \neg Y$

$(A \Rightarrow B)\,,\ (X \Rightarrow B)\,,\ X\,,\ \neg A\,,\ \neg Y$

$(X \Rightarrow B)\,,\ X\,,\ \neg A\,,\ \neg Y$

$B\,,\ X\,,\ \neg A\,,\ \neg Y$

$\bigcirc$

FIG. 36 – Tableau sémantique simplifié : $D_2 \not\models C_2$

On introduit les propositions a, b, c, d, e pour modéliser la présence éventuelle de A, B, C, D, E, respectivement. Le coffre peut être ouvert dans toute situation vérifiant la formule

$$\Phi =_{def} (a \wedge b) \vee (a \wedge c) \vee (b \wedge d \wedge e)\,.$$

En utilisant la règle de distributivité de la disjonction sur la conjonction, on voit que Φ est logiquement équivalente à la conjonction des 12 clauses suivantes :

$$\begin{array}{lll}
(a \vee a \vee b)\,, & (a \vee a \vee d)\,, & (a \vee a \vee e)\,,\\
(b \vee a \vee b)\,, & (b \vee a \vee d)\,, & (b \vee a \vee e)\,,\\
(a \vee c \vee b)\,, & (a \vee c \vee d)\,, & (a \vee c \vee e)\,,\\
(b \vee c \vee b)\,, & (b \vee c \vee d)\,, & (b \vee c \vee e)\,.
\end{array}$$

On peut omettre les répétitions de littéraux au sein d'une clause, ce qui simplifie les clauses en

$$\begin{array}{lll}
(a \vee b)\,, & (a \vee d)\,, & (a \vee e)\,,\\
(a \vee b)\,, & (b \vee a \vee d)\,, & (b \vee a \vee e)\,,\\
(a \vee c \vee b)\,, & (a \vee c \vee d)\,, & (a \vee c \vee e)\,,\\
(b \vee c)\,, & (b \vee c \vee d)\,, & (b \vee c \vee e)\,.
\end{array}$$

De plus, si une clause (vue comme un ensemble de littéraux) en contient une autre, la clause *contenante* peut être omise ; seules quatre clauses subsistent :

$$1 : (a \vee b)\,,\ 2 : (a \vee d)\,,\ 3 : (a \vee e)\,,\ 4 : (b \vee c)\,.$$

Ceci montre qu'une solution à quatre serrures $(1, 2, 3, 4)$ convient, avec la distribution de clés suivante :

$$A : 1,2,3\,;\ \ B : 1,4\,;\ \ C : 4\,;\ \ D : 2\,;\ \ E : 3\,.$$

Penser ou payer ! Vous entrez dans un pub écossais et le barman vous dit : “Vous voyez ces trois hommes ? L’un d’eux est Monsieur X, qui dit toujours la vérité, un autre est Monsieur Y, qui ment toujours, et le troisième est Monsieur Z, qui répond au hasard sans écouter les questions. Vous pouvez poser trois questions (appelant une réponse par oui ou non), en indiquant chaque fois lequel des trois doit répondre. Si après cela vous pouvez identifier correctement ces messieurs, ils vous offrent un whisky !”. Comment vous y prenez-vous ?

Il est clair que les réponses de Monsieur Z sont sans intérêt, aussi une bonne tactique consistera à repérer d’abord quelqu’un qui n’est pas Monsieur Z ; cela peut se faire au moyen d’une question bien choisie. C’est à ce quelqu’un que l’on posera les deux dernières questions.
On pose à l’un des hommes (I) la question

Votre voisin de gauche (G) est-il plus menteur que votre voisin de droite (D) ?

Examinons, pour les six dispositions possibles, la réponse fournie, étant entendu que Y est plus menteur que Z, lui-même plus menteur que X :

I	G	D	Réponse exacte	Réponse fournie
X	Y	Z	*oui*	*oui*
X	Z	Y	*non*	*non*
Y	X	Z	*non*	*oui*
Y	Z	X	*oui*	*non*
Z	X	Y	*non*	*oui/non*
Z	Y	X	*oui*	*oui/non*

On observe que si la réponse fournie est *oui*, le voisin de gauche n’est jamais Z ; c’est donc à lui que l’on s’adressera pour les deux questions suivantes. de même, si la réponse fournie est *non*, c’est au voisin de droite, qui n’est jamais Z, que l’on s’adressera.

Dans les deux cas, on posera ensuite une question dont on connaît la réponse, par exemple “Etes-vous Monsieur Z ?”, qui identifiera ce nouvel interlocuteur (X si “non”, Y si “oui”). La troisième question, “Votre voisin de gauche est-il Monsieur Z ?”, permettra de compléter les identifications.

L’enquête policière. Cinq suspects (A, B, C, D et E) sont interrogés à propos d’un crime. Voici leurs déclarations :

A : C et D mentent.
B : A et E mentent.
C : B et D mentent.
D : C et E mentent.
E : A et B mentent.

Que peut-on en déduire ?

Si x signifie “X dit la vérité”, on peut modéliser les déclarations en les formules

A : $a \equiv (\neg c \wedge \neg d)$
B : $b \equiv (\neg a \wedge \neg e)$
C : $c \equiv (\neg b \wedge \neg d)$
D : $d \equiv (\neg c \wedge \neg e)$
E : $e \equiv (\neg a \wedge \neg b)$

Remarque. On pourrait imaginer d’autres interprétations des déclarations, et donc d’autres modélisations.

Supposons que A dise la vérité ; ceux qui disent le contraire ont donc menti et on en déduit

A : $\neg c, \neg d$
B : $\neg b$
C : $c \equiv \neg d$
D : $d \equiv \neg c$
E : $\neg e$

On obtient une contradiction, ce qui montre que A a menti. La situation est donc

A : $\neg a$, $c \vee d$
B : $b \equiv \neg e$
C : $c \equiv (\neg b \wedge \neg d)$
D : $d \equiv (\neg c \wedge \neg e)$
E : $e \equiv \neg b$

Si B dit la vérité, on obtient $\neg a$, $c \vee d$, b, $\neg e$, $\neg c$, d.

Si B a menti, on obtient $\neg a$, $c \vee d$, $\neg b$, e, c, $\neg d$

En conclusion, A est certainement menteur mais, pour les quatre autres suspects, il y a deux possibilités : B et D disent la vérité et C et E mentent, ou B et D mentent et C et E disent la vérité.

3.6 La méthode de résolution

La méthode de résolution est la technique de preuve la plus souvent utilisée dans les programmes de démonstration automatique ; elle intervient souvent, sous une forme ou sous une autre, dans les programmes d’intelligence artificielle. Cette méthode opère par réfutation : elle permet de démontrer la validité de A en démontrant l’inconsistance de $\neg A$. Plus généralement, pour démontrer $E \models A$, on prouve l’inconsistance de $E \cup \{\neg A\}$.

On peut appliquer la méthode de résolution à n’importe quel ensemble de formules mais, comme pour les méthodes vues antérieurement, il est plus simple de se limiter à un certain type de formules, appelées formes clausales.

3.6.1 Formes normales

Formes normales en algèbre. L’expression $(x^2 - 4x)(x + 3) + (2x - 1)^2 + 4x - 19$ est un polynôme, mais ses propriétés ne sont pas directement apparentes. On souhaitera donc *normaliser* le polynôme, c’est-à-dire trouver un polynôme équivalent (en fait, égal) mais de forme plus “agréable”. Le type de *forme normale* ou de *forme canonique* choisi dépendra

des propriétés que l'on souhaite mettre en évidence et/ou utiliser, et aussi de l'existence et de l'efficacité de l'algorithme de recherche de la forme choisie. On a notamment les formes suivantes :

$x^3 + 3x^2 - 12x - 18$ (somme de monômes de degrés décroissants) ;

$(x - 3)(x + 3 - \sqrt{3})(x + 3 + \sqrt{3})$ (produit de facteurs du premier degré) ;

$[(x + 3)x - 12]x - 18$ (forme de Horner).

Les formules propositionnelles, comme les polynômes, peuvent prendre diverses formes normales ; nous en introduisons ici deux.

Forme normale disjonctive. La figure 37 montre qu'une formule est toujours équivalente à une disjonction de conjonctions de littéraux.

p	q	r	$p \Rightarrow q$	$(p \Rightarrow q) \Rightarrow r$
V	**V**	**V**	**V**	**V**
V	**V**	**F**	**V**	**F**
V	**F**	**V**	**F**	**V**
V	**F**	**F**	**F**	**V**
F	**V**	**V**	**V**	**V**
F	**V**	**F**	**V**	**F**
F	**F**	**V**	**V**	**V**
F	**F**	**F**	**V**	**F**

$$
\begin{array}{ll}
 & (\ p \land \ \ q \land \ \ r) \\
\lor & (\ p \land \neg q \land \ \ r) \\
\lor & (\ p \land \neg q \land \neg r) \\
\lor & (\neg p \land \ \ q \land \ \ r) \\
\lor & (\neg p \land \neg q \land \ \ r)
\end{array}
$$

FIG. 37 – Table de vérité et forme normale disjonctive.

On appelle *forme normale disjonctive* (FND) toute disjonction de conjonctions de littéraux. Toute formule est donc équivalente à une FND.

Remarque. Il s'agit de disjonctions et de conjonctions *généralisées*, c'est-à-dire à nombre quelconque (mais fini) de termes.[61]

Un *cube* est une conjonction de littéraux, c'est-à-dire une formule $(\ell_1 \land \ell_2 \land \cdots \land \ell_n)$, $(n \in \mathbb{N})$, où les ℓ_i sont des littéraux. On écrit parfois $\bigwedge\{\ell_1, \ldots, \ell_n\}$, ou $\bigwedge_i \ell_i$, ou simplement $\{\ell_1, \ldots, \ell_n\}$.

Remarque. Dans ce contexte, les connecteurs "0-aires" $true$ et $false$ ne sont pas utilisés.

On observe immédiatement qu'un cube est inconsistant si et seulement s'il contient une paire complémentaire de littéraux ; de plus, le cube vide est le seul cube valide.

Une forme normale disjonctive est inconsistante si et seulement si tous ses cubes sont inconsistants. En particulier, la forme normale disjonctive vide est inconsistante.

[61] NB : $\bigvee\emptyset \leftrightarrow false$, $\bigwedge\emptyset \leftrightarrow true$, $\bigvee\{A\} \leftrightarrow A \leftrightarrow \bigwedge\{A\}$, $\bigvee\{A, B\} \leftrightarrow A \lor B$, $\bigwedge\{A, B\} \leftrightarrow A \land B$.

Forme normale conjonctive. Une *clause* est une disjonction de littéraux. Une telle formule est parfois représentée par la notation ensembliste $\bigvee\{\ell_i : i = 1, \ldots, n\}$, voire $\{\ell_i : i = 1, \ldots, n\}$.[62]

Remarque. Selon le contexte, la notation $p\overline{q}r$ peut représenter le cube $p \land \neg q \land r$ ou la clause $p \lor \neg q \lor r$. Cette notation compacte mais ambiguë est à éviter.

La seule clause inconsistante est la *clause vide*, représentée par **F** ou par $\Box$. (On n'utilise pas les connecteurs *true* et *false*.) Une clause est valide si et seulement si elle comporte une paire complémentaire de littéraux. Une *clause unitaire* est une clause composée d'un seul littéral.

Une *forme normale conjonctive* (FNC, ou CNF pour *Conjunctive Normal Form*) est une conjonction de clauses, c'est-à-dire une conjonction de disjonctions de littéraux. Voici un exemple et un contre-exemple :

– $(\neg p \lor q \lor r) \land (\neg q \lor r) \land (\neg r)$ – CNF
– $(\neg p \lor q \lor r) \land \neg(\neg q \lor r) \land (\neg r)$ – non CNF

Une forme normale conjonctive est valide si et seulement toutes ses clauses sont valides. En particulier, la forme normale conjonctive vide est valide. Toute formule du calcul des propositions peut être transformée en une forme normale conjonctive équivalente.

Rappel. Les clauses, cubes et formes normales sont des formules ; on les traite souvent comme des ensembles (conjonctifs ou disjonctifs) mais ces ensembles sont toujours *finis*.

Intérêt des formes normales. Une forme normale doit être, idéalement

– assez générale pour que chaque formule soit réductible à une forme normale équivalente,
– aussi restrictive que possible, pour que les algorithmes qui traitent les formes normales soient plus simples que les algorithmes généraux.

Des formes normales conjonctives ou disjonctives distinctes peuvent être équivalentes.

Exemple. La forme normale disjonctive

$(p \land q \land r) \lor (p \land \neg q \land r) \lor (p \land \neg q \land \neg r) \lor (\neg p \land q \land r) \lor (\neg p \land \neg q \land \neg r)$

se simplifie en

$(p \land r) \lor (\neg q \land \neg r) \lor (q \land r)$

Algorithme de normalisation. On ne considère que la forme normale *conjonctive*.

1. Eliminer les connecteurs autres que $\neg$, $\lor$, $\land$.
2. Utiliser les lois de De Morgan pour propager les occurrences de $\neg$ vers l'intérieur.

$$
\begin{array}{rcl}
\neg(A \land B) & \leftrightarrow & (\neg A \lor \neg B) \\
\neg(A \lor B) & \leftrightarrow & (\neg A \land \neg B)
\end{array}
$$

3. Eliminer les doubles négations.

$$\neg\neg A \leftrightarrow A$$

4. Utiliser les lois de distributivité de $\lor$ par rapport à $\land$ pour éliminer $\land$ des disjonctions.

$$
\begin{array}{rcl}
A \lor (B \land C) & \leftrightarrow & (A \lor B) \land (A \lor C) \\
(A \land B) \lor C & \leftrightarrow & (A \lor C) \land (B \lor C)
\end{array}
$$

[62] Il faut se méfier de cette dernière notation : une clause est un ensemble *disjonctif* de littéraux.

Une formule en forme conjonctive normale équivaut à un *ensemble* (conjonctif) de clauses ; on dit aussi *forme clausale*.
Exercice. Comment obtenir une forme disjonctive normale équivalente à A au départ d'une forme conjonctive normale équivalente à $\neg A$?

Exemple de normalisation. On récrit la formule $(\neg p \Rightarrow \neg q) \Rightarrow (p \Rightarrow q)$ en forme conjonctive normale.

$(\neg p \Rightarrow \neg q) \Rightarrow (p \Rightarrow q)$	
$\neg(\neg\neg p \vee \neg q) \vee (\neg p \vee q)$	(élimination $\Rightarrow$)
$(\neg\neg\neg p \wedge \neg\neg q) \vee (\neg p \vee q)$	(propagation $\neg$)
$(\neg p \wedge q) \vee (\neg p \vee q)$	(double négation)
$(\neg p \vee \neg p \vee q) \wedge (q \vee \neg p \vee q)$	(distributivité)

Une forme normale conjonctive de $(\neg p \Rightarrow \neg q) \Rightarrow (p \Rightarrow q)$ est $(\neg p \vee \neg p \vee q) \wedge (q \vee \neg p \vee q)$.
Une forme plus simple est $\neg p \vee q$.

Algorithme de normalisation (variante). Une variante intéressante de l'algorithme de normalisation est représentée à la figure 38. La donnée manipulée est un ensemble conjonctif L de disjonctions généralisées. Initialement, l'unique élément de L est la formule donnée A (vue comme une disjonction généralisée à un terme). La valeur finale de L est une FNC équivalente à A. On appelle *non-clause* toute disjonction (généralisée) dont au moins un terme n'est pas un littéral. La preuve de terminaison est analogue à celle pour la construction des tableaux sémantiques. La valeur finale de L est l'ensemble de clauses cherché.
Remarque. Cet algorithme n'est qu'une reformulation de l'algorithme précédent. L'intérêt est lié à la méthode de résolution vue plus loin.
Exemple de normalisation (variante)

1.	$\{(\neg p \Rightarrow \neg q) \Rightarrow (p \Rightarrow q)\}$	*Init*
2.	$\{\neg(\neg p \Rightarrow \neg q) \vee (p \Rightarrow q)\}$	$\beta, 1$
3.	$\{\neg(\neg p \Rightarrow \neg q) \vee \neg p \vee q\}$	$\beta, 2$
4.	$\{\neg p \vee \neg p \vee q\, ,\ \neg\neg q \vee \neg p \vee q\}$	$\alpha, 3$
5.	$\{\neg p \vee \neg p \vee q\, ,\ q \vee \neg p \vee q\}$	$\alpha, 4$

La forme normale requise est donc $(\neg p \vee \neg p \vee q) \wedge (q \vee \neg p \vee q)$.
Elle se simplifie en $(\neg p \vee q) \wedge (\neg p \vee q)$, et puis en $\neg p \vee q$.

Simplifications des formes clausales. Les formes clausales ou ensembles conjonctifs de clauses fournis par l'algorithme de normalisation peuvent souvent être simplifiés.

1. On peut supprimer les répétitions de littéraux au sein d'une même clause.
   Exemple : $(\neg p \vee q \vee \neg p) \wedge (r \vee \neg p) \leftrightarrow (\neg p \vee q) \wedge (r \vee \neg p)$.

$L := \{A\}$;
Tant que L comporte une non-clause faire
  $\{\bigwedge L \leftrightarrow A$ est invariant$\}$
  choisir une non-clause $D \in L$;
  choisir un non-littéral $t \in D$;
  $*$ si $t = \neg\neg t'$ faire
    $D' := (D - t) + t'$;
    $\{D \leftrightarrow D'\}$
    $L := (L \backslash \{D\}) \cup \{D'\}$
  $*$ si $t = \alpha$ faire
    $t_1 := \alpha_1;\ t_2 := \alpha_2;$
    $D_1 := (D - t) + t_1;\ D_2 := (D - t) + t_2;$
    $\{D \leftrightarrow D_1 \wedge D_2\}$
    $L := (L \backslash \{D\}) \cup \{D_1, D_2\}$
  $*$ si $t = \beta$ faire
    $t_1 := \beta_1;\ t_2 := \beta_2;$
    $D' := ((D - t) + t_1) + t_2;$
    $\{D \leftrightarrow D'\}$
    $L := (L \backslash \{D\}) \cup \{D'\}$

FIG. 38 – Algorithme de normalisation.

2. Les clauses valides (elles contiennent une paire complémentaire de littéraux) peuvent être supprimées.
   Exemple : $(\neg p \vee q \vee p) \wedge (r \vee \neg p) \leftrightarrow (r \vee \neg p)$.
3. Une clause *contenant* une autre clause peut être supprimée.
   Exercice : justifier la règle.
   Exemple : $(r \vee q \vee \neg p) \wedge (\neg p \vee r) \leftrightarrow (\neg p \vee r)$.

Ces simplifications élémentaires sont faciles à mettre en œuvre mais ne conduisent pas à une forme normale unique. Par exemple, elles ne permettent pas de réduire $(p \vee \neg q) \wedge q$ en $p \wedge q$.

3.6.2 La règle de résolution

Définition. On sait qu'un ensemble de clauses S est inconsistant si et seulement si $S \models \Box$. ($\Box$ est la clause vide qui dénote *false*.) Cela suggère de montrer l'inconsistance de S en essayant de dériver $\Box$ (*false*) à partir de S, au moyen d'un mécanisme adéquat.

Soient A, B, X des formules et soit v une interprétation. Supposons $v(A \vee X) = \mathbf{V}$ et $v(B \vee \neg X) = \mathbf{V}$. Si $v(X) = \mathbf{V}$, alors $v(B) = \mathbf{V}$ et donc $v(A \vee B) = \mathbf{V}$. Si $v(X) = \mathbf{F}$, alors $v(A) = \mathbf{V}$ et donc $v(A \vee B) = \mathbf{V}$. En conclusion, $\{(A \vee X), (B \vee \neg X)\} \models (A \vee B)$.

Cette règle très simple est appelée *règle de résolution* dans le cas où X est une proposition et où A, B sont des clauses. Avec la notation habituelle, on l'écrit

$$\frac{A \vee X\,, \quad B \vee \neg X}{A \vee B}$$

Fermeture par résolution. On définit par induction la relation $\vdash_{\mathcal{R}}$ (que nous noterons simplement $\vdash$) entre un ensemble de clauses et une clause ; c'est la plus petite relation vérifiant les deux conditions suivantes :

1. Si $C \in S$, alors $S \vdash C$.
2. Soient $C_1 = (C'_1 \vee p)$ et $C_2 = (C'_2 \vee \neg p)$; si $S \vdash C_1$ et $S \vdash C_2$, alors $S \vdash C'_1 \vee C'_2$.

Les deux clauses C_1 et C_2 sont dites *résolvables* (par rapport à p) ; leur *résolvante* est la clause $res(C_1, C_2) =_{def} C'_1 \vee C'_2$.
Si S est un ensemble de clauses, S^R dénote la *fermeture de S par résolution*, c'est-à-dire le plus petit sur-ensemble de S contenant les résolvantes de ses éléments. On a $S^R = \{C : S \vdash C\} = \{C : S^R \vdash C\}$.

Adéquation de la règle de résolution. Soient S un ensemble de clauses et C une clause. On doit montrer que si $S \vdash C$, alors $S \models C$. Il suffit de montrer que la relation $\models$ (restreinte aux ensembles de clauses et aux clauses) vérifie les deux conditions définissant la relation $\vdash_{\mathcal{R}}$:

1. Si $C \in S$, alors $S \models C$.
2. Soient $C_1 = (C'_1 \vee p)$ et $C_2 = (C'_2 \vee \neg p)$; si $S \models C_1$ et $S \models C_2$, alors $S \models C'_1 \vee C'_2$.

La première condition est évidente, la seconde est une conséquence de l'énoncé $\{(A \vee X), (B \vee \neg X)\} \models (A \vee B)$, valable quelles que soient les formules A, B et X.

Remarque. On déduit de ceci que les ensembles S et S^R sont logiquement équivalents, pour tout ensemble S de clauses.

3.6.3 Complétude de la méthode de résolution

Introduction. Si S est un ensemble de clauses, si A est une clause et si $S \models A$, a-t-on nécessairement $S \vdash_{\mathcal{R}} A$? La réponse est clairement négative : on a

$$\{p, \neg p\} \models q\,,$$

mais on n'a pas

$$\{p, \neg p\} \vdash_{\mathcal{R}} q\,.$$

Cela n'est pas gênant, dans la mesure où on ne cherchera pas à utiliser la résolution pour établir directement $S \models A$, mais plutôt $S, \neg A \models \Box$. On démontrera et utilisera le théorème suivant.

Théorème. Si $S \models \Box$, alors $S \vdash_{\mathcal{R}} \Box$.

Cette "complétude affaiblie" (cas particulier où la clause à dériver est toujours la clause vide) est aussi puissante que la complétude usuelle (non satisfaite ici) puisqu'on peut toujours se ramener au cas particulier. C'est pourquoi la "complétude affaiblie" est nommée simplement "complétude".

Arbre sémantique. Soient S une formule ou un ensemble (conjonctif) de formules et $(p_1, p_2, \ldots)$ une énumération de Π_S, l'ensemble (fini ou dénombrable) des propositions présentes dans S. L'*arbre sémantique* de S est un arbre binaire complet dont toutes les branches de gauche de niveau i sont étiquetées par p_i et toutes les branches de droite de niveau i sont étiquetées par $\neg p_i$.

L'arbre sémantique de S décrit toutes les interprétations possibles de S. Chaque chemin $\mathcal{C}$ dans l'arbre allant de la racine à un nœud n de niveau i définit

– un ensemble de propositions, soit $\Pi(n) = \{p_1, \ldots, p_i\}$;
– une interprétation pour cet ensemble de propositions, soit v_n ; on a $v_n(p_k) = \mathbf{V}$ si $p_k \in \mathcal{C}$ et $v_n(p_k) = \mathbf{F}$ si $\neg p_k \in \mathcal{C}$.

Exemple. Soit $S = \{p \vee q, p \vee r, \neg q \vee \neg r, \neg p\}$, un ensemble de clauses. Le lexique Π_S est $\{p, q, r\}$. Un arbre sémantique relatif à ce lexique est donné à la figure 39. L'arbre est fini puisque Π_S est fini. Comme S est inconsistant, chaque feuille peut être étiquetée par une clause fausse pour l'interprétation correspondante.

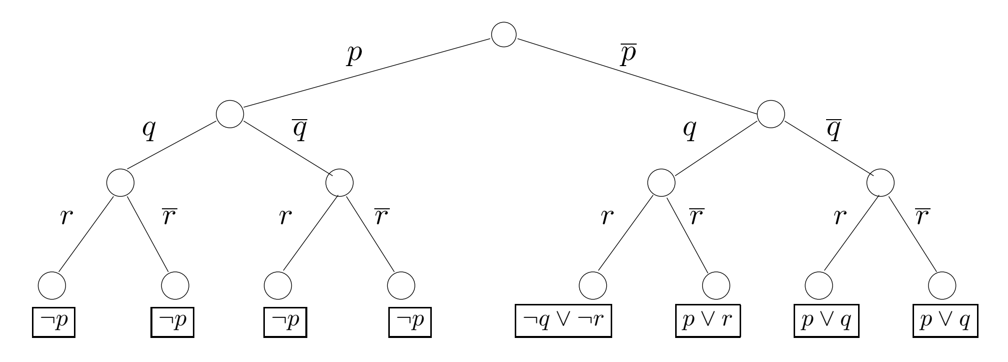


FIG. 39 – Arbre sémantique montrant l'inconsistance d'un ensemble de clauses.

Preuve de complétude dans le cas fini. Soit S inconsistant et fini ; on doit prouver $S \vdash \Box$. Comme d'habitude, on souhaite une preuve constructive, c'est-à-dire un moyen effectif d'obtenir $\Box$ au départ de S, par applications répétées de la règle de résolution.

Soit $\mathcal{A}$ un arbre sémantique pour S. Chaque chemin dans cet arbre allant de la racine à un nœud n définit un ensemble de propositions $\Pi(n)$ et une interprétation v_n pour cet ensemble ; v_n rend vrais les littéraux étiquetant le chemin.

S est inconsistant, donc S est falsifié par toutes les interprétations définies par les feuilles de $\mathcal{A}$. Par conséquent, à chaque feuille f de l'arbre correspond au moins une clause $C_f \in S$ telle que

$$\Pi_{C_f} \subseteq \Pi(f) = \Pi_S \quad \text{et} \quad v_f(C_f) = \mathbf{F}\,.$$

(Π_{C_f} est l'ensemble des propositions intervenant dans C_f.) La feuille f est étiquetée C_f.

Soit $S^R = S \cup \{C : S \vdash C\}$. On va montrer qu'il est possible d'étiqueter chaque nœud intérieur n de l'arbre au moyen d'une clause $C_n \in S^R$ telle que

$$\Pi_{C_n} \subseteq \Pi(n) \subseteq \Pi_S \quad \text{et} \quad v_n(C_n) = \mathbf{F}.$$

De cette manière, on aura au noeud racine r une clause $C_r \in S^R$ telle que

$$\Pi_{C_r} \subseteq \Pi(r) \quad \text{et} \quad v_r(C_r) = \mathbf{F}.$$

Or comme $\Pi(r) = \emptyset$ et $v_r(C_r) = \mathbf{F}$, on aura nécessairement $C_r = \Box$.

L'étiquetage se fait en remontant des feuilles vers la racine.
Soit une paire de nœuds n_1, n_2 de l'arbre ayant un nœud père n commun tel que

$$\Pi(n_1) = \Pi(n_2) = \Pi(n) \cup \{p\}.$$

On suppose
$C_{n_1} \in S^R$ et $\Pi_{C_{n_1}} \subseteq \Pi(n_1)$ et $v_{n_1}(C_{n_1}) = \mathbf{F}$
$C_{n_2} \in S^R$ et $\Pi_{C_{n_2}} \subseteq \Pi(n_2)$ et $v_{n_2}(C_{n_2}) = \mathbf{F}$
L'étiquette C_n de n sera C_{n_1} ou C_{n_2} ou $res(C_{n_1}, C_{n_2})$ et ne contiendra ni p ni $\neg p$; cela suggère la politique de choix suivante :

- Si $p \notin \Pi_{C_{n_i}}$ pour $i = 1$ ou 2, alors $C_n = C_{n_i}$.
- Si $p \in \Pi_{C_{n_1}}$ et $p \in \Pi_{C_{n_2}}$:
  $v_{n_1}(C_{n_1}) = \mathbf{F}$ implique $C_{n_1} = C'_{n_1} \vee \neg p$ et $v_{n_2}(C_{n_2}) = \mathbf{F}$ implique $C_{n_2} = C'_{n_2} \vee p$.
  On pose $C_n = C'_{n_1} \vee C'_{n_2} = res_p(C_{n_1}, C_{n_2})$.

Dans les deux cas, on a $C_n \in S^R$ et $\Pi_{C_n} \subseteq \Pi(n)$ et $v_n(C_n) = \mathbf{F}$.
Ceci achève la démonstration du cas fini.
Un exemple d'étiquetage complet de l'arbre sémantique est donné à la figure 40.

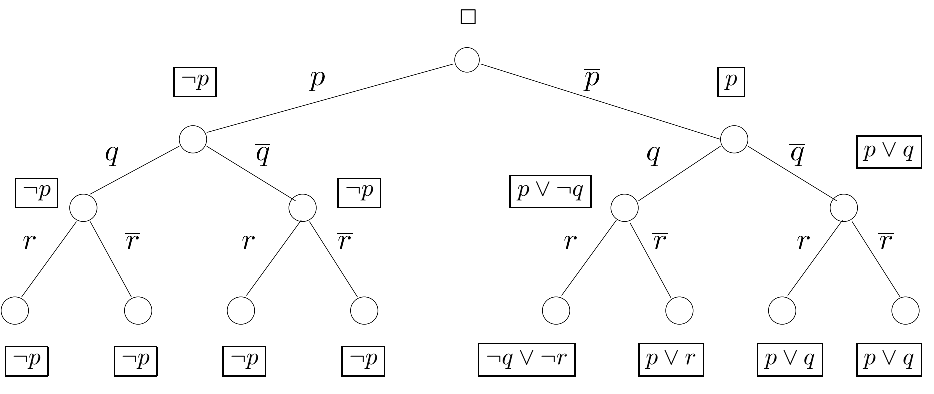


FIG. 40 – Arbre sémantique montrant l'inconsistance d'un ensemble de clauses.

Complétude dans le cas infini, preuve indirecte. On a jusqu'ici supposé que S était fini mais, vu le théorème de compacité, le résultat

$$\Box \in S^R \quad \text{ssi} \quad S \text{ est inconsistant}$$

reste valable si S est infini. En effet, si S est inconsistant, il admet un sous-ensemble fini S_f inconsistant ; on en déduit $\Box \in S_f^R$, d'où a fortiori $\Box \in S^R$.

Complétude dans le cas infini, preuve directe. Prouver directement ce résultat revient à donner une autre preuve, moins abstraite, du théorème de compacité. On se limite au cas habituel où le lexique Π est un ensemble dénombrable. L'arbre sémantique correspondant $\mathcal{A}$ comporte une infinité de branches, elles-mêmes infinies. A chaque nœud n on associe v_n comme précédemment ; le nœud n est un *nœud-échec* s'il existe $C_n \in S$ telle que $v_n(C_n) = \mathbf{F}$.

On obtient l'arbre $\mathcal{B}$ en élaguant $\mathcal{A}$, de telle sorte que les feuilles de $\mathcal{B}$ soient des nœuds-échecs et que les nœuds intérieurs ne le soient pas. (Les feuilles de $\mathcal{B}$ ne sont pas nécessairement toutes au même niveau.)

L'ensemble S étant inconsistant, toutes les branches de $\mathcal{B}$ sont finies. Un arbre binaire dont toutes les branches sont finies est nécessairement fini (c'est un cas particulier du classique lemme de König, dont nous (re)verrons la démonstration au paragraphe suivant). On applique à $\mathcal{B}$ la technique d'étiquetage introduite pour le cas fini. L'ensemble $S_0 \subset S$ des clauses associées aux feuilles de $\mathcal{B}$ est donc tel que $\Box \in S_0^R$, et S_0 est un sous-ensemble fini inconsistant de S. On a donc prouvé que tout ensemble inconsistant de clauses construit au moyen d'un lexique dénombrable admet un sous-ensemble fini inconsistant. Toute formule étant logiquement équivalente à un ensemble (conjonctif) de clauses, on a en fait démontré que tout ensemble inconsistant de formules construit au moyen d'un lexique dénombrable admet un sous-ensemble fini inconsistant.

Lemme de König. *Définition.* Un arbre *fini* est un arbre comportant un nombre fini de nœuds. Un arbre est *finitaire* si chaque nœud a un nombre fini de fils.

Lemme. Tout arbre infini finitaire a au moins une branche infinie.
Démonstration. Considérons un arbre infini finitaire. Soit n_0 sa racine. L'arbre est infini, donc n_0 a un nombre infini de descendants. L'arbre est finitaire, donc n_0 a un descendant direct, soit n_1, qui a un nombre infini de descendants. De même, n_1 doit avoir un descendant direct, soit n_2, qui a un nombre infini de descendants. On peut itérer indéfiniment ; on obtient ainsi la branche infinie $n_0, n_1, n_2, \ldots$

3.6.4 Procédure de résolution

Si S est un ensemble de clauses, on note $\mathcal{M}_S$ l'ensemble des modèles de S. L'ensemble S est inconsistant si et seulement si $\mathcal{M}_S = \emptyset$. L'algorithme représenté à la figure 41 met en œuvre la vérification d'inconsistance par résolution.
Remarque sur l'invariant de boucle. Ajouter à S des conséquences logiques de ses éléments ne change pas l'ensemble $\mathcal{M}_S$ des modèles de S.

$S := S_0$; (S_0 est un ensemble de clauses)
$\{\mathcal{M}_S = \mathcal{M}_{S_0}\}$
Tant que $\Box \notin S$, répéter :
  choisir $p \in \Pi_S$,
    $C_1 = (C'_1 \vee p) \in S$,
    $C_2 = (C'_2 \vee \neg p) \in S$;
  $S := S \cup \{res(C_1, C_2)\}$
  $\{\mathcal{M}_S = \mathcal{M}_{S_0}\}$

FIG. 41 – Procédure de résolution.

Remarque sur la procédure de choix. On admet qu'aucune paire de clauses résolvables ne peut être choisie plus d'une fois ; cela garantit la *terminaison* puisqu'un lexique de n propositions donne lieu à 3^n clauses distinctes non valides. Le programme peut se terminer normalement (garde falsifiée) ou anormalement (plus de choix possible).
Terminaison normale. Si la garde devient fausse, la valeur finale S_f vérifie $\mathcal{M}_{S_f} = \mathcal{M}_{S_0}$ et $\Box \in S_f$, ce qui implique l'*inconsistance* de S_f et de S_0.
Terminaison anormale. Si toutes les résolvantes ont été calculées sans produire $\Box$, on a $\mathcal{M}_{S_f} = \mathcal{M}_{S_0}$ et $\Box \notin S_f$. Cela implique la *consistance* de S_f et de S_0.

Remarque. Une dérivation de $\Box$ (*false*) à partir de S est appelée une *réfutation* de S.

Exemples de réfutations. Soit S l'ensemble des quatre clauses suivantes :

1. $p \vee q$
2. $p \vee r$
3. $\neg q \vee \neg r$
4. $\neg p$

Cet ensemble est inconsistant ; il admet au moins une réfutation. Comme souvent, il en existe plusieurs, telles que

5.	q	$(1,4)$
6.	r	$(2,4)$
7.	$\neg q$	$(3,6)$
8.	$\Box$	$(5,7)$

5.	$p \vee \neg r$	$(1,3)$
6.	q	$(1,4)$
7.	$p \vee \neg q$	$(2,3)$
8.	r	$(2,4)$
9.	p	$(2,5)$
10.	$\neg r$	$(3,6)$
11.	$\neg q$	$(3,8)$
12.	$\neg r$	$(4,5)$
13.	$\neg q$	$(4,7)$
14.	$\Box$	$(4,9)$

Soit S' l'ensemble des deux clauses suivantes :

1. p
2. $\neg p \vee q$

La seule dérivation possible est

3. q $(1,2)$

qui ne produit pas la clause vide. L'ensemble est donc consistant.

Soit S'' l'ensemble des trois clauses suivantes :

1. p
2. $\neg p \vee q$
3. $\neg q$

On obtient immédiatement la réfutation

4. q $(1,2)$
5. $\Box$ $(3,4)$

qui prouve l'inconsistance de l'ensemble.

Efficacité de la résolution. On sait que l'algorithme de résolution est correct et se termine toujours, mais est-il efficace ? Même si on évite de produire plusieurs fois la même résolvante, il est clair que le nombre de clauses produites peut être exponentiel en la taille de S (ou en la taille du lexique Π_S).

La plupart du temps, l'emploi d'une stratégie adaptée permet d'obtenir une efficacité acceptable (dans le cas où l'ensemble de départ est inconsistant). On peut cependant construire des "cas pathologiques" pour lesquels aucune stratégie efficace n'existe.

3.7 Exercice de récapitulation

Soit A la formule $[(p \wedge q) \vee (r \Rightarrow s)] \Rightarrow [(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))]$.
On utilise diverses méthodes pour prouver la validité de A.

3.7.1 Méthode directe

Elle consiste à considérer toutes les interprétations. La formule A comporte quatre propositions distinctes, il y a donc $2^4 = 16$ interprétations. Il est plus commode de structurer l'analyse que de procéder en 16 étapes ; on utilise aussi les règles de simplifications élémentaires concernant $a \circ b$, où $\circ$ est un connecteur binaire. Ces règles s'appliquent dès que a et b sont égaux ou opposés, et aussi si l'un des opérandes est **V** ou **F**. On simplifie aussi les doubles négations.

$$\begin{array}{l}
p = \mathbf{V} : \ [(T \wedge q) \vee (r \Rightarrow s)] \Rightarrow [(T \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))], \\
\quad [q \vee (r \Rightarrow s)] \ \Rightarrow \ [T \wedge (q \vee (r \Rightarrow s))], \\
\quad [q \vee (r \Rightarrow s)] \ \Rightarrow \ [q \vee (r \Rightarrow s)], \\
\quad \mathbf{V}\,; \\
p = \mathbf{F} : \ [(F \wedge q) \vee (r \Rightarrow s)] \Rightarrow [(F \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))], \\
\quad [F \vee (r \Rightarrow s)] \ \Rightarrow \ [(r \Rightarrow s) \wedge (q \vee (r \Rightarrow s))], \\
\quad (r \Rightarrow s) \ \Rightarrow \ [(r \Rightarrow s) \wedge (q \vee (r \Rightarrow s))]\,; \\
\quad q = \mathbf{V} : \ (r \Rightarrow s) \ \Rightarrow \ [(r \Rightarrow s) \wedge (T \vee (r \Rightarrow s))], \\
\quad\quad (r \Rightarrow s) \ \Rightarrow \ [(r \Rightarrow s) \wedge T], \\
\quad\quad (r \Rightarrow s) \ \Rightarrow \ (r \Rightarrow s), \\
\quad\quad \mathbf{V}\,; \\
\quad q = \mathbf{F} : \ (r \Rightarrow s) \ \Rightarrow \ [(r \Rightarrow s) \wedge (F \vee (r \Rightarrow s))], \\
\quad\quad (r \Rightarrow s) \ \Rightarrow \ [(r \Rightarrow s) \wedge (r \Rightarrow s)], \\
\quad\quad (r \Rightarrow s) \ \Rightarrow \ (r \Rightarrow s), \\
\quad\quad \mathbf{V}.
\end{array}$$

Diverses variantes existent selon le nombre de règles simplificatrices admises (assez réduit ici) et le niveau auquel on les applique (ici, tous).

3.7.2 Méthode algébrique

Elle consiste à utiliser les lois algébriques pour simplifier les formules. Cette méthode est très efficace si on fait les meilleurs choix … mais elle est très inefficace sinon !

$$[(p \wedge q) \vee (r \Rightarrow s)] \ \Rightarrow \ [(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))]$$

{ Distributivité de $\wedge$ sur $\vee$ (antécédent) }

$$[(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))] \Rightarrow [(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))]$$

{ $X \ \Rightarrow \ X \ \leftrightarrow \ true$ pour tout X }

$$true$$

3.7.3 Tableau sémantique (notation réduite)

$$\neg A,\ (p \wedge q) \vee (r \Rightarrow s),$$
$$\neg[(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))]$$

↙ ↘

$\neg(p \vee (r \Rightarrow s))$, $\neg p$, $\neg(r \Rightarrow s)$ — ↙ ↘ — $p \wedge q,\ p,\ q$ × ; $r \Rightarrow s$ ×

$\neg(q \vee (r \Rightarrow s))$, $\neg q$, $\neg(r \Rightarrow s)$ — ↙ ↘ — $p \wedge q,\ p,\ q$ × ; $r \Rightarrow s$ ×

On observe une certaine redondance dans les calculs ; c'est le prix à payer pour une méthode facilement mécanisable.

3.7.4 Réduction à la forme conjonctive

On applique les règles habituelles :

$$[(p \wedge q) \vee (r \Rightarrow s)] \Rightarrow [(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))]$$
$$\neg[(p \wedge q) \vee \neg r \vee s] \vee [(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s)]$$
$$[(\neg p \vee \neg q) \wedge r \wedge \neg s] \vee [(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s)]$$
$$(\neg p \wedge r \wedge \neg s) \vee (\neg q \wedge r \wedge \neg s) \vee [(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s)]$$

Une conjonction est valide si et seulement si tous ses termes sont valides. On prouve donc séparément la validité des deux formules

$A_1 \ =_{def} \ (\neg p \wedge r \wedge \neg s) \vee (\neg q \wedge r \wedge \neg s) \vee p \vee \neg r \vee s$

et

$A_2 \ =_{def} \ (\neg p \wedge r \wedge \neg s) \vee (\neg q \wedge r \wedge \neg s) \vee q \vee \neg r \vee s.$

Chacune de ces formules se réduit à une conjonction de 9 clauses ; on considère seulement la formule A_1. Les clauses sont

$\neg p \vee \neg q \vee p \vee \neg r \vee s$
$\neg p \vee r \vee p \vee \neg r \vee s$
$\neg p \vee \neg s \vee p \vee \neg r \vee s$
$r \vee \neg q \vee p \vee \neg r \vee s$
$r \vee r \vee p \vee \neg r \vee s$
$r \vee \neg s \vee p \vee \neg r \vee s$
$\neg s \vee \neg q \vee p \vee \neg r \vee s$
$\neg s \vee r \vee p \vee \neg r \vee s$
$\neg s \vee \neg s \vee p \vee \neg r \vee s$

Chaque clause comporte une paire complémentaire de littéraux et est donc valide.

3.7.5 Résolution

On commence par réduire $\neg A$ en forme clausale.

$$\neg([(p \wedge q) \vee (r \Rightarrow s)] \Rightarrow [(p \vee (r \Rightarrow s)) \wedge (q \vee (r \Rightarrow s))])$$
$$[(p \wedge q) \vee \neg r \vee s] \wedge \neg[(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s)]$$
$$(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s) \wedge [(\neg p \wedge r \wedge \neg s) \vee (\neg q \wedge r \wedge \neg s)]$$
$$\ldots$$

Cela donne 11 clauses :

1. $p \vee \neg r \vee s$
2. $q \vee \neg r \vee s$
3. $\neg p \vee \neg q$
4. $\neg p \vee r$
5. $\neg p \vee \neg s$
6. $r \vee \neg q$
7. r
8. $r \vee \neg s$
9. $\neg s \vee \neg q$
10. $\neg s \vee r$
11. $\neg s$

On déduit la clause vide $\square$ par résolution :

12. $p \vee s$ $\quad 1, 7$
13. $q \vee s$ $\quad 2, 7$
14. p $\quad 11, 12$
15. q $\quad 11, 13$
16. $\neg q$ $\quad 3, 14$
17. $\square$ $\quad 15, 16$

Remarque. Les clauses valides ou sur-ensembles d'autres clauses sont inutiles et peuvent être supprimées d'emblée. Dans le cas présent, toutes les clauses minimales (1, 2, 3, 7 et 11) ont été utilisées.

3.7.6 Résolution généralisée

La déduction
$X \vee Y\,,\ \neg X \vee Z\ \models\ Y \vee Z$
reste correcte si X n'est pas un littéral.
On utilise aussi les trois règles suivantes :

$\alpha \vee X\ \models\ \alpha_1 \vee X\,,$
$\alpha \vee X\ \models\ \alpha_2 \vee X\,,$
$\beta \vee X\ \models\ \beta_1 \vee \beta_2 \vee X\,.$

On obtient alors une réfutation plus courte, sans réduction préalable à la forme clausale :

1. $\neg A$
2. $(p \wedge q) \vee \neg r \vee s$ $\quad 1, \alpha_1$
3. $\neg[(p \vee \neg r \vee s) \wedge (q \vee \neg r \vee s)]$ $\quad 1, \alpha_2$
4. $\neg(p \vee \neg r \vee s) \vee \neg(q \vee \neg r \vee s)$ $\quad 3, \beta$
5. $p \vee \neg r \vee s$ $\quad 2, \alpha_1$
6. $q \vee \neg r \vee s$ $\quad 2, \alpha_2$
7. $\neg(q \vee \neg r \vee s)$ $\quad 4, 5, R$
8. $\square$ $\quad 6, 7, R$

3.7.7 Méthode *ad-hoc*

Elle consiste à tirer parti des particularités de la formule étudiée … c'est-à-dire à n'avoir pas de méthode !

On peut observer, par exemple, que

$$[(p \wedge q)\ \vee\ (r \Rightarrow s)] \Rightarrow [(p \vee (r \Rightarrow s))\ \wedge\ (q \vee (r \Rightarrow s))]$$

est de la forme

$$(A\ \vee\ B)\ \Rightarrow\ (C\ \wedge\ D)\,,$$

et qu'une telle formule est valide si et seulement si les formules $A \Rightarrow C$, $A \Rightarrow D$, $B \Rightarrow C$, $B \Rightarrow D$ sont valides. On doit donc prouver

$$\models\ (p \wedge q)\ \Rightarrow\ (p \vee (r \Rightarrow s))\,,$$
$$\models\ (p \wedge q)\ \Rightarrow\ (q \vee (r \Rightarrow s))\,,$$
$$\models\ (r \Rightarrow s)\ \Rightarrow\ (p \vee (r \Rightarrow s))\,,$$
$$\models\ (r \Rightarrow s)\ \Rightarrow\ (q \vee (r \Rightarrow s))\,,$$

ce qui est évident dans chaque cas.

4 Méthodes déductives : le système de Hilbert

4.1 Introduction

Nous avons vu qu'une théorie est l'ensemble des conséquences logiques d'un ensemble donné de formules, appelées *axiomes* ou *postulats*. Ces conséquences logiques sont appelées *théorèmes*. Développer une théorie consiste donc à repérer les théorèmes parmi les formules construites au moyen du lexique (du langage) utilisé pour introduire les postulats. Deux grandes techniques existent pour cela, la méthode analytique et la méthode synthétique.

Jusqu'ici, nous avons utilisé la méthode analytique, sous la forme d'une procédure de décision. Pour analyser une formule propositionnelle, il suffit de construire un ou deux tableau(x) sémantique(s). Cette approche est excellente … quand elle est possible. En logique prédicative, qui est la logique des mathématiciens, on ne dispose pas en général d'une procédure de décision ; même quand elle existe, elle peut être difficile à mettre en œuvre. De plus, une procédure de décision pour la validité ne donne guère d'information sur le lien sémantique entre axiomes et théorèmes. Enfin, les procédures de décision sont souvent inefficaces parce qu'elles ne réutilisent pas les résultats. On ne peut pas, en général, accélérer la validation d'un théorème sur base d'autres théorèmes antérieurement démontrés.

Les mathématiciens utilisent le plus souvent la méthode synthétique. Des théorèmes simples sont obtenus à partir des postulats au moyen de quelques mécanismes de raisonnement. Ces mêmes mécanismes, appliqués aux théorèmes simples, permettent de démontrer des théorèmes plus difficiles, et ainsi de suite. L'approche synthétique est également utilisée dans les autres sciences exactes, et notamment en physique ; dans une certaine mesure, on utilise également l'approche synthétique en médecine, en psychologie, en sociologie, etc. Un aspect typique de cette approche est l'exploitation de résultats antérieurs pour produire des résultats nouveaux. Le principal avantage de cette approche est sa généralité. La méthode synthétique s'accommode d'un ensemble infini de postulats (chaque preuve n'en utilise qu'un nombre fini) ; elle permet d'isoler les postulats nécessaires à la production d'un théorème donné, ce qui permet notamment de déterminer si un théorème subsiste ou non quand l'ensemble des postulats est modifié. La théorie est développée de manière modulaire, chaque théorème pouvant être assimilé à un postulat supplémentaire, disponible pour l'obtention de nouveaux théorèmes.

Ces avantages ont un prix. L'obtention de théorèmes par l'approche synthétique est un processus foncièrement non déterministe, pouvant requérir créativité, inventivité … et tâtonnement, au contraire de l'approche analytique dans laquelle le non-déterminisme est inexistant (tables de vérité) ou peu important (tableaux sémantiques). Des choix inadéquats conduisent à des théorèmes corrects mais inintéressants ; on voit qu'une certaine forme de créativité est nécessaire ici. On peut même dire que le talent du mathématicien consiste essentiellement à opérer les bons choix, ceux qui conduisent à valider (ou à infirmer) les conjectures les plus remarquables.[63]

La logique propositionnelle est suffisamment élémentaire pour être correctement appréhendée au moyen des seules méthodes analytiques. Nous introduisons néanmoins l'approche synthétique pour préparer le lecteur à son utilisation dans le cadre plus complexe de la logique prédicative.

[63]Une autre facette du talent du mathématicien est l'aptitude à créer de nouveaux ensembles de postulats conduisant à des théories intéressantes.

4.2 Axiomes et règle d'inférence

Le système formel $\mathcal{H}$ est constitué

– de *schémas d'axiomes* ; on a

1. $\vdash A \Rightarrow (B \Rightarrow A)$
2. $\vdash (A \Rightarrow (B \Rightarrow C)) \;\Rightarrow\; ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$
3. $\vdash (\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B)$

– de la *règle du Modus ponens (MP)* :

$$\frac{\vdash A \qquad \vdash A \Rightarrow B}{\vdash B}$$

A, B et C sont des formules quelconques, n'utilisant que les connecteurs "$\neg$" et "$\Rightarrow$".[64] Un axiome proprement dit s'obtient en *instanciant* l'un des trois schémas, c'est-à-dire en remplaçant A, B, C par des formules. Par exemple, la formule $(p \Rightarrow q) \Rightarrow (p \Rightarrow (p \Rightarrow q))$ est un axiome, obtenu en appliquant la substitution $[A/(p \Rightarrow q),\ B/p]$ au premier schéma. La notation "$\vdash \varphi$" se lit "φ est un théorème". Rappelons ici que tout axiome est un théorème. Le système de Hilbert comporte la seule règle d'inférence "Modus ponens". Cette règle permet d'obtenir le théorème B au départ des théorèmes A et $A \Rightarrow B$.

4.3 Preuves

Une *preuve* dans $\mathcal{H}$ est une séquence de formules, chaque formule étant

– l'instance d'un axiome, ou
– inférée de deux formules la précédant dans la séquence, au moyen de la règle d'inférence Modus ponens.

Par définition, tout élément d'une preuve, et en particulier le dernier, est un théorème ; si A est le dernier élément de la séquence, celle-ci est une *preuve de A*. A titre d'exemple, une preuve du conditionnel $p \Rightarrow p$ est donnée à la figure 42.

Remarques. Par facilité, chaque élément d'une preuve est précédé d'un numéro d'ordre et suivi d'une brève justification ; "Axiome 1" veut dire "instance du schéma d'axiome 1" et "4, 3, MP" veut dire "obtenu à partir des formules de numéros 4 et 3 (prémisses) par la règle du Modus ponens". La preuve donnée ici établit que la formule $(p \Rightarrow p)$ est un théorème, ou encore que l'assertion $\vdash (p \Rightarrow p)$ (lire : "$(p \Rightarrow p)$ est un théorème") est un *métathéorème* (c'est-à-dire un théorème au sens mathématique courant ; le préfixe "méta" est souvent omis).[65] L'expression

[64]Il existe des variantes permettant l'emploi de tous les connecteurs habituels, mais la version présentée ici est plus simple, sans être réellement restrictive ; on considère $p \vee q$ comme une abréviation de $\neg p \Rightarrow q$, et $p \wedge q$ comme une abréviation de $\neg(p \Rightarrow \neg q)$.

[65]Signalons que $(p \Rightarrow p)$ est une formule, donc un objet du langage, tandis que $\vdash (p \Rightarrow p)$ est une assertion, donc un objet du métalangage.

1. $\vdash p \Rightarrow ((p \Rightarrow p) \Rightarrow p)$	(Axiome 1)
2. $\vdash (p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p))$	(Axiome 2)
3. $\vdash (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$	(1, 2, MP)
4. $\vdash p \Rightarrow (p \Rightarrow p)$	(Axiome 1)
5. $\vdash p \Rightarrow p$	(4, 3, MP)

FIG. 42 – Exemple de preuve dans le système de Hilbert.

$(A \Rightarrow A)$ est un *schéma de théorème* ; toute instance d'un schéma de théorème est un théorème. On transforme facilement la preuve de $(p \Rightarrow p)$ en une preuve de $(p \Rightarrow q) \Rightarrow (p \Rightarrow q)$, par exemple.

La preuve donnée plus haut peut se représenter de manière arborescente (figure 43). Cette représentation est plus naturelle que la représentation séquentielle introduite plus haut, mais n'est guère utilisée à cause de son encombrement.

$$\dfrac{\dfrac{\vdash p \Rightarrow ((p \Rightarrow p) \Rightarrow p) \qquad \vdash (p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p))}{\vdash (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)} \qquad \vdash p \Rightarrow (p \Rightarrow p)}{\vdash p \Rightarrow p}$$

FIG. 43 – Représentation arborescente d'une preuve.

On peut aussi (figure 44) ne mentionner dans l'arborescence que les numéros des formules impliquées, ce qui réduit l'encombrement.

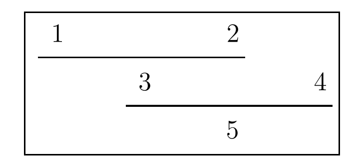


FIG. 44 – Représentation arborescente abrégée d'une preuve.

Il existe une nette analogie entre une preuve de Hilbert (représentée de manière arborescente) et une dérivation de séquent, mais il y a aussi quelques différences :

- Le symbole "$\vdash$" remplace le symbole "$\rightarrow$".
- Les antécédents sont vides (pour l'instant).
- Les succédents comportent un seul élément.
- Le sens de "axiome" a changé.
- L'unique règle est le Modus ponens, qui n'est pas analytique, ni réversible.

En dépit de ces différences, on peut dire que le système de Hilbert est un calcul de séquent (synthétique).

Remarque. Le mot "axiome" a en fait plusieurs sens relativement voisins, mais qu'il convient de distinguer. Dans le cadre de la méthode (analytique) des séquents, un axiome est un séquent d'un type particulier, dont la validité est immédiate. Dans le cadre du système de Hilbert, un axiome est une tautologie d'un certain type, obtenue par instantiation d'un schéma particulier. Dans les deux cas, l'idée de validité est importante. En revanche, dans le langage courant, dans le langage mathématique général et dans le cadre plus technique des théories logiques (surtout prédicatives), les axiomes ne sont pas des tautologies mais des énoncés consistants dont on souhaite étudier l'ensemble des conséquences logiques. Dans le cadre de cette étude seulement, les axiomes sont considérés comme toujours vrais ; il s'agit donc d'une validité "locale", limitée à un certain contexte. En pratique, ce contexte peut paraître universel. C'est le cas de l'énoncé "l'addition est commutative", traduit par la formule $\forall x \, \forall y \, (x{+}y = y{+}x)$. Cette formule n'est pourtant valide que si on interprète de manière adéquate le symbole fonctionnel "$+$" et le symbole prédicatif "$=$".

Notons aussi que le mot "*postulat*" est synonyme du mot "*axiome*" mais que ce dernier insiste plus sur l'aspect "vérité universelle" tandis que "*postulat*" met plus l'accent sur l'aspect relatif de la validité. Le célèbre énoncé d'Euclide "Par tout point extérieur à une droite passe une et une seule parallèle à cette droite" est un axiome de la géométrie classique[66] et un postulat — auquel il est parfois profitable de renoncer — de la géométrie moderne. De même, la commutativité de l'addition est un axiome (ou un théorème) en arithmétique et un postulat en théorie des groupes.

4.4 Dérivations

Un ensemble U de formules quelconques étant donné, une *dérivation* ou *preuve avec hypothèses* dans $\mathcal{H}$ est une séquence de formules, chaque formule étant

- une *hypothèse* (élément de U), ou
- une instance d'un axiome, ou
- inférée de deux formules précédentes, au moyen du Modus ponens.

Le métathéorème relatif à une preuve avec hypothèses s'écrit $U \vdash_{\mathcal{H}} A$ ou $U \vdash A$. Un exemple de dérivation est donné à la figure 45. Comme pour les preuves, les lignes sont numérotées et accompagnées d'une courte justification. En outre, l'ensemble des hypothèses est rappelé à chaque ligne. On verra plus loin pourquoi.

Remarque. En général, le dernier élément A d'une dérivation n'est pas un théorème. On verra plus loin que $U \vdash A$ a lieu si et seulement si on a $U \models A$; en particulier, $\vdash A$ a lieu si et seulement si A est une tautologie.

Remarques. Il est souvent plus facile d'établir $A, B, C \vdash D$ que $\vdash A \Rightarrow (B \Rightarrow (C \Rightarrow D))$, mais on montrera qu'une dérivation du premier énoncé se convertit automatiquement en une preuve du second ; la dérivation ci-dessus établit donc indirectement

$$\vdash (p \Rightarrow (q \Rightarrow r)) \Rightarrow (q \Rightarrow (p \Rightarrow r)) .$$

[66] Après en avoir longtemps été une conjecture, dont les mathématiciens ont finalement déterminé qu'elle ne pouvait être déduite des autres axiomes.

1.	$p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash p \Rightarrow (q \Rightarrow r)$	(Hypothèse)
2.	$p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash p$	(Hypothèse)
3.	$p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash q \Rightarrow r$	(1, 2, MP)
4.	$p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash q$	(Hypothèse)
5.	$p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash r$	(3, 4, MP)

FIG. 45 – Exemple de dérivation dans le système de Hilbert.

La représentation arborescente, style séquent, reste possible. Les hypothèses deviennent les éléments des antécédents. (Le sens du mot "hypothèse" a donc changé.)

4.5 Quelques résultats utiles

Nous voyons ici trois résultats permettant de raccourcir les preuves ou, plus exactement, d'écrire des "textes" qui ne sont pas, à strictement parler, des preuves, mais des argumentations (souvent plus courtes que les preuves elles-mêmes) établissant l'existence d'une preuve d'un théorème donné.

4.5.1 Principes de composition et de substitution uniforme

Théorème. Tout théorème peut être utilisé comme un axiome dans une preuve.[67]
Démonstration. On obtient une preuve (au sens strict) en remplaçant chaque théorème utilisé comme un axiome par une preuve de ce théorème.

Théorème. Si C est un théorème et si $p_1, \dots, p_n$ sont des propositions deux à deux distinctes, alors $C(p_1/A_1, \dots, p_n/A_n)$ est un théorème.
Démonstration. L'application d'une substitution uniforme à toutes les lignes d'une preuve produit toujours une preuve.

4.5.2 Règles d'inférence dérivées

L'écriture $\dfrac{U_1 \vdash A_1\,, \cdots, U_n \vdash A_n}{U \vdash B}$
est une *règle d'inférence dérivée* ; elle exprime que s'il existe des dérivations pour chacune des n prémisses, alors il existe une dérivation pour la conclusion. Une règle est *adéquate* ou *correcte* si elle exprime une vérité.

Remarque. Une règle dérivée est correcte ... si l'on peut s'en passer, c'est-à-dire si toute dérivation l'utilisant peut être convertie en une dérivation ne l'utilisant pas. Une fois que nous

[67] On notera la différence de sens entre les deux emplois du mot "théorème" ; la première occurrence aurait pu être "métathéorème".

aurons montré que $U \vdash A$ est assimilable à $U \models A$, on pourra prouver facilement qu'une règle est correcte. Par exemple, la règle

$$\frac{U\,, \neg X \vdash X}{U \vdash X}$$

est correcte, parce que si X est conséquence logique de $U \cup \{\neg X\}$, alors X est est conséquence logique de U. Nous devrons cependant établir directement certaines règles, nécessaires pour démontrer que les relations $\vdash$ et $\models$ sont coextensives.
Remarque. Dans ce contexte, "U, A" abrège "$U \cup \{A\}$".

On notera que les principes de composition et de substitution uniforme peuvent se traduire par des règles dérivées, de même que le principe de monotonie, selon lequel une hypothèse supplémentaire n'altère pas les dérivations faites sans elle. On a

$$\frac{U \vdash A \qquad U, A \vdash B}{U \vdash B}$$

$$\frac{U \vdash A}{U[p/B] \vdash A[p/B]}$$

$$\frac{U \vdash A}{U, B \vdash A}$$

4.6 Règle de déduction

Cette règle dérivée essentielle s'écrit $\dfrac{U, A \vdash B}{U \vdash A \Rightarrow B}$.

Remarque. On voit que cette règle provient directement d'une règle (réversible et de type α) du calcul des séquents. Elle formalise une démarche courante en mathématiques :

- on doit démontrer $A \Rightarrow B$;
- on suppose A ;
- on en dérive B.

On a déjà vu l'utilité potentielle de la règle (pour peu que son adéquation soit établie) :

- $p \Rightarrow (q \Rightarrow r),\, q,\, p \vdash r$ est trivial ;
- $\vdash (p \Rightarrow (q \Rightarrow r)) \Rightarrow (q \Rightarrow (p \Rightarrow r))$ ne l'est pas.

Notons enfin que la règle est réversible, la règle inverse étant une simple variante du Modus Ponens.

4.6.1 Adéquation de la règle de déduction

On doit démontrer que toute preuve utilisant la règle de déduction peut se récrire en une preuve ne l'utilisant pas. Cela revient à transformer, étape par étape, une preuve de $U, A \vdash B$ en une preuve de $U \vdash A \Rightarrow B$.

Démonstration. On suppose donnée une preuve Π_1 de $U, A \vdash B$, dont chaque étape est du type $U, A \vdash X$. On construit une preuve Π_2 de $U \vdash (A \Rightarrow B)$ en remplaçant chaque étape de Π_1

$n.\ U, A \vdash X$

par une ou plusieurs nouvelles étapes dont la dernière est

n'. $U \vdash A \Rightarrow X$.

On distingue quatre cas :

1. X est *un axiome* ;
2. X est *une hypothèse de l'ensemble* U ;
3. X est *la nouvelle hypothèse* A ;
4. X est *inféré par Modus ponens*.

– **Dans les cas 1 et 2,** l'étape
  n. $U, A \vdash X$ (Ai ou H)
  est remplacée par les trois étapes suivantes :
  $n'-2$. $U \vdash X$ (Ai ou H)
  $n'-1$. $U \vdash X \Rightarrow (A \Rightarrow X)$ (A1)
  n'. $U \vdash A \Rightarrow X$ ($n'-2$, $n'-1$, MP)
– **Dans le cas 3,** l'étape $U, A \vdash A$ est remplacée par cinq étapes calquées sur la démonstration de $\vdash (p \Rightarrow p)$ donnée plus haut ; la dernière de ces cinq nouvelles étapes est naturellement $U \vdash (A \Rightarrow A)$.
– **Dans le cas 4,** la preuve Π_1 comporte les étapes
  i. $U, A \vdash Y$ (...)
  j. $U, A \vdash Y \Rightarrow X$ (...)
  n. $U, A \vdash X$ (i, j, MP)
  Le préfixe déjà construit de Π_2 comportera
  i'. $U \vdash A \Rightarrow Y$ (...)
  j'. $U \vdash A \Rightarrow (Y \Rightarrow X)$ (...)
  Le fragment de Π_2 relatif à la $n$ième étape de Π_1 sera :
  $n'-2$. $U \vdash (A \Rightarrow (Y \Rightarrow X)) \Rightarrow ((A \Rightarrow Y) \Rightarrow (A \Rightarrow X))$ (A2)
  $n'-1$. $U \vdash (A \Rightarrow Y) \Rightarrow (A \Rightarrow X)$ (j', $n'-2$, MP)
  n'. $U \vdash (A \Rightarrow X)$ (i', $n'-1$, MP)

Ceci achève la démonstration.

4.7 Théorèmes et règles dérivées supplémentaires

Dans le système de Hilbert, les preuves sont très faciles à *vérifier* mais nettement plus difficiles à *construire* ; cette situation est habituelle avec les méthodes synthétiques. Nous donnons ici quelques théorèmes utiles et quelques règles dérivées supplémentaires.

4.7.1 Théorèmes supplémentaires

Nous donnons ici neuf théorèmes importants :

1. $\vdash (A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$
2. $\vdash (A \Rightarrow (B \Rightarrow C)) \Rightarrow (B \Rightarrow (A \Rightarrow C))$
3. $\vdash \neg A \Rightarrow (A \Rightarrow B)$
4. $\vdash A \Rightarrow (\neg A \Rightarrow B)$
5. $\vdash \neg\neg A \Rightarrow A$
6. $\vdash A \Rightarrow \neg\neg A$
7. $\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$
8. $\vdash B \Rightarrow (\neg C \Rightarrow \neg(B \Rightarrow C))$
9. $\vdash (B \Rightarrow A) \Rightarrow ((\neg B \Rightarrow A) \Rightarrow A)$

Il s'agit en fait de schémas de théorèmes ; chacune des lettres A, B et C désigne ici n'importe quelle formule.

Dans la suite, on évitera la *construction* des preuves ; on préférera démontrer simplement l'*existence* d'une preuve. Les notions de *dérivation* et de *règle dérivée* ont pour but de faciliter ces démonstrations d'existence.

A titre d'exemple, une dérivation du théorème 6 est donnée à la figure 46, les théorèmes 1 à 5 étant supposés déjà démontrés. On voit ici l'utilité capitale du principe de composition ; la règle de déduction facilite le travail de justification ... qui n'est quand même pas évident.

1. $A, \neg\neg\neg A \vdash \neg\neg\neg A \Rightarrow \neg A$	(Composition, th. 5)
2. $A, \neg\neg\neg A \vdash \neg\neg\neg A$	(Hypothèse)
3. $A, \neg\neg\neg A \vdash \neg A$	(1, 2, MP)
4. $A \vdash \neg\neg\neg A \Rightarrow \neg A$	(Déduction, 3)
5. $A \vdash (\neg\neg\neg A \Rightarrow \neg A) \Rightarrow (A \Rightarrow \neg\neg A)$	(Axiome 3)
6. $A \vdash A \Rightarrow \neg\neg A$	(4, 5, MP)
7. $A \vdash A$	(Hypothèse)
8. $A \vdash \neg\neg A$	(6, 7, MP)
9. $\vdash A \Rightarrow \neg\neg A$	(Déduction, 8)

FIG. 46 – Justification de $\vdash A \Rightarrow \neg\neg A$.

Remarque. Une règle dérivée évidente, le plus souvent utilisée implicitement, est la *règle d'augmentation* ; elle s'écrit

$$\frac{U \vdash A}{U, B \vdash A}$$

et exprime qu'une hypothèse disponible ne doit pas nécessairement être utilisée.

4.7.2 Quelques autres règles dérivées

Notons d'emblée un puissant moyen de construire des règles dérivées.

Métarègle. Si $\vdash A \Rightarrow B$, alors $\dfrac{U \vdash A}{U \vdash B}$ est une règle dérivée correcte.

Cela formalise une démarche intuitive :

1. Ayant démontré A en supposant U, soit $U \vdash A$,
2. on utilise le théorème $\vdash A \Rightarrow B$
3. et on applique la règle du Modus ponens à (1) et (2) pour obtenir $U \vdash B$.

On obtient ainsi diverses règles dérivées utiles, dont voici quatre exemples :

$$\frac{U \vdash \neg B \Rightarrow \neg A}{U \vdash A \Rightarrow B} \qquad \textit{Contraposée}$$

$$\frac{U \vdash A \Rightarrow B \qquad U \vdash B \Rightarrow C}{U \vdash A \Rightarrow C} \qquad \textit{Transitivité}$$

$$\frac{U \vdash \neg\neg A}{U \vdash A} \qquad \textit{Double négation}$$

$$\frac{U \vdash A \Rightarrow (B \Rightarrow C)}{U \vdash B \Rightarrow (A \Rightarrow C)} \qquad \textit{Echange d'antécédents}$$

La règle de *contraposition* formalise le raisonnement par l'absurde. La règle de *transitivité* formalise l'emploi de lemmes "en cascade" : pour démontrer $\vdash A \Rightarrow B$, on démontre les lemmes $\vdash A \Rightarrow C_1, \vdash C_1 \Rightarrow C_2, \ldots, \vdash C_n \Rightarrow B$. Par application répétée de la règle de transitivité, on déduit $\vdash A \Rightarrow B$. La règle d'*échange d'antécédents* indique que l'ordre dans lequel des hypothèses sont faites n'est pas important. La règle de la double négation est couramment utilisée en mathématique. Elle n'est pas innocente dans le cadre de la philosophie, de la linguistique, de l'informatique. Par exemple, la phrase

> Il n'est pas vrai que je suis mécontent.

n'est pas tout à fait équivalente à

> Je suis content.

De même, un programme qui ne produit pas deux valeurs $x \neq y$, n'est pas nécessairement un programme qui produit deux valeurs $x = y$.

Signalons encore la règle de disjonction des cas, qui s'écrit

$$\frac{U, B \vdash A \qquad U, \neg B \vdash A}{U \vdash A}$$

Voici un schéma de démonstration de cette règle importante.

$U, B \vdash A$	Prémisse
$U \vdash B \Rightarrow A$	Déduction
$\vdash (B \Rightarrow A) \Rightarrow ((\neg B \Rightarrow A) \Rightarrow A)$	Théorème 9, § 4.7.1
$U \vdash (\neg B \Rightarrow A) \Rightarrow A$	MP
$U, \neg B \vdash A$	Prémisse
$U \vdash \neg B \Rightarrow A$	Déduction
$U \vdash A$	MP

Remarque. Un peu de créativité ... ou de patience est nécessaire pour démontrer le théorème utilisé dans le développement ci-dessus.

4.8 Adéquation et complétude du système de Hilbert

Le système de Hilbert est un mécanisme permettant d'engendrer des théorèmes ; on souhaite naturellement que tout théorème soit une tautologie (adéquation) et que toute tautologie soit un théorème (complétude). On peut établir un résultat plus général : si A est une formule et si U est un ensemble de formules, alors

$$U \models A$$

est vrai si et seulement si

$$U \vdash A$$

est vrai. (Seuls le conditionnel et la négation sont admis.)

4.8.1 Adéquation du système de Hilbert

Montrer que tous les théorèmes sont des tautologies revient à montrer que tous les éléments d'une preuve Π sont des tautologies. C'est une conséquence immédiate des lemmes suivants, eux-mêmes évidents.

Lemme. Toute instance des schémas d'axiomes est une tautologie.

Lemme. Si A et $A \Rightarrow B$ sont des tautologies, alors B est une tautologie.

On démontre de même que si l'assertion $U \vdash A$ apparaît dans une dérivation, alors la formule A est conséquence logique de l'ensemble U.

4.8.2 Lemme de Kalmar

Le système de Hilbert est, nous venons de le prouver, un mécanisme de production de tautologies. Il faut aussi prouver que ce système produit *toutes* les tautologies. Le point de vue constructif adopté dans ce texte nous incite donc à élaborer une stratégie d'utilisation du système de Hilbert, permettant de construire à la demande une preuve dont le dernier élément est une tautologie donnée.

Le système le plus simple de vérification de tautologie est sans doute la méthode des tables de vérité.[68] Le lemme de Kalmar spécifie qu'à chaque ligne d'une table de vérité correspond une dérivation dans le système de Hilbert.

Lemme. Soit A une formule construite à partir des propositions $p_1, \ldots, p_n$ et des connecteurs "$\neg$" et "$\Rightarrow$". Soit v une interprétation. Si on définit p'_k comme p_k ou $\neg p_k$ selon que $v(p_k)$ est **V** ou **F**, et si on définit A' comme A ou $\neg A$ selon que $v(A)$ est **V** ou **F**, on a

$$\{p'_1, \ldots, p'_n\} \vdash A'$$

Exemple. Du fragment de table de vérité

p	q	r	s	$(p \Rightarrow q) \Rightarrow \neg(\neg r \Rightarrow s)$
F	**F**	**V**	**V**	**F**

le lemme de Kalmar permet de déduire

$$\{\neg p, \neg q, r, s\} \;\vdash\; \neg[(p \Rightarrow q) \Rightarrow \neg(\neg r \Rightarrow s)]\,.$$

Remarque. Le lemme de Kalmar permet de "coder" une ligne de table de vérité dans le système de Hilbert ; il contribue donc à établir que $U \models A$ implique $U \vdash A$.

4.8.3 Démonstration du lemme de Kalmar

Remarque préliminaire. On va utiliser les lemmes suivants :

$C \;\vdash\; B \Rightarrow C\,,$

$B \;\vdash\; \neg\neg B\,,$

$\neg B \;\vdash\; B \Rightarrow C\,,$

$\neg C,\, B \;\vdash\; \neg(B \Rightarrow C)\,.$

dont les démonstrations sont évidentes si l'on tient compte respectivement de l'axiome 1 (§ 4.2) et des théorèmes 6, 3 et 8 (§ 4.7.1)

Démonstration. On raisonne par induction sur la structure de la formule A.

Cas de base : la formule A est une proposition p_k.
Si $v(p_k) = \mathbf{V}$, l'énoncé se réduit à $\{\ldots, p_k, \ldots\} \vdash p_k$.
Si $v(p_k) = \mathbf{F}$, l'énoncé se réduit à $\{\ldots, \neg p_k, \ldots\} \vdash \neg p_k$.

Premier cas inductif : la formule A est la négation $\neg B$.

[68]Ne confondons pas "le plus simple" et "le plus efficace".

Premier sous-cas.

$v(B) = \mathbf{F}$ et $v(A) = \mathbf{V}$.

$\{p'_1, \ldots, p'_n\} \vdash B'\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg B\,,$

$\{p'_1, \ldots, p'_n\} \vdash A\,,$

$\{p'_1, \ldots, p'_n\} \vdash A'\,.$

Deuxième sous-cas.

$v(B) = \mathbf{V}$ et $v(A) = \mathbf{F}$.

$\{p'_1, \ldots, p'_n\} \vdash B'$

$\{p'_1, \ldots, p'_n\} \vdash B\,,$

$B \vdash \neg\neg B\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg\neg B\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg A\,,$

$\{p'_1, \ldots, p'_n\} \vdash A'\,.$

Second cas inductif : la formule A est le conditionnel $B \Rightarrow C$.

Premier sous-cas.

$v(C) = \mathbf{V}$ et $v(A) = \mathbf{V}$.

$\{p'_1, \ldots, p'_n\} \vdash C'\,,$

$\{p'_1, \ldots, p'_n\} \vdash C\,,$

$C \vdash (B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash (B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash A\,,$

$\{p'_1, \ldots, p'_n\} \vdash A'\,.$

Deuxième sous-cas.

$v(B) = \mathbf{F}$ et $v(A) = \mathbf{V}$.

$\{p'_1, \ldots, p'_n\} \vdash B'\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg B\,,$

$\neg B \vdash (B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash (B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash A\,,$

$\{p'_1, \ldots, p'_n\} \vdash A'\,.$

Troisième sous-cas.

$v(B) = \mathbf{V}$, $v(C) = \mathbf{F}$ et $v(A) = \mathbf{F}$.

$\{p'_1, \ldots, p'_n\} \vdash B'\,,$

$\{p'_1, \ldots, p'_n\} \vdash B\,,$

$\{p'_1, \ldots, p'_n\} \vdash C'\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg C\,,$

$\neg C,\, B \;\vdash\; \neg(B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg(B \Rightarrow C)\,,$

$\{p'_1, \ldots, p'_n\} \vdash \neg A\,,$

$\{p'_1, \ldots, p'_n\} \vdash A'\,.$

Remarque. On a raisonné par cas pour établir $\{p'_1, \ldots, p'_n\} \vdash A'$, sans pour autant utiliser la règle dérivée de disjonction des cas.[69] En fait, cette règle dérivée est, comme les autres, une version particulière et formelle d'une technique de raisonnement (informel).

4.8.4 Complétude du système de Hilbert

Soit A une tautologie construite à partir des propositions $p_1, \ldots, p_n$ et des connecteurs "$\neg$" et "$\Rightarrow$". D'après le lemme de Kalmar, on a

$$\{p'_1, \ldots, p'_n\} \vdash A\,,$$

où p'_k est indifféremment p_k ou $\neg p_k$.
(On a toujours $A' = A$.)

De ces 2^n théorèmes on tire, par disjonction des cas sur p'_n, les 2^{n-1} théorèmes suivants :

$$\{p'_1, \ldots, p'_{n-1}\} \vdash A\,,$$

[69]En revanche, cette règle sera utilisée explicitement au paragraphe suivant.

et, plus généralement, les 2^k théorèmes suivants, pour tout $k \in \{0, 1, \ldots, n\}$:

$$\{p'_1, \ldots, p'_k\} \vdash A\,,$$

et donc, en particulier ($k = 0$) le théorème

$$\vdash A\,,$$

ce qui achève la démonstration.

5 Logique prédicative : syntaxe et sémantique

5.1 Introduction

Nous avons vu d'emblée que la principale limitation de la logique propositionnelle est l'impossibilité de modéliser adéquatement des propositions "paramétriques", dont la valeur de vérité dépend de la signification de termes contenus dans la proposition. Reconsidérons les exemples suivants :

– Les entiers naturels x et $x + 2$ sont premiers.
– Il existe un naturel x tel que x et $x + 2$ sont premiers.
– Il existe une infinité de nombres premiers x tels que $x + 2$ est aussi premier.
– Il fera beau à tel endroit, à tel instant.
– Il fera beau à Liège le 29 avril de l'an 2021.
– $x^2 + y^2 = z^2$.
– $x^3 + y^3 = z^3$.
– Il existe des entiers strictement positifs x, y, z tels que $x^2 + y^2 = z^2$.
– Il existe des entiers strictement positifs x, y, z tels que $x^3 + y^3 = z^3$.

On observe d'abord que, faute de pouvoir analyser des propositions paramétriques telles que "Les entiers naturels x et $x + 2$ sont premiers" et "$x^2 + y^2 = z^2$", on ne peut pas non plus analyser des propositions non paramétriques telles que "Il existe un naturel x tel que x et $x+2$ sont premiers" et "Il existe des entiers strictement positifs x, y, z tels que $x^2 + y^2 = z^2$". En effet, il est fréquent que des propositions non paramétriques admettent comme composants (dans un sens à préciser) des propositions paramétriques. Dans la mesure où l'approche compositionnelle semble incontournable, on voit que la logique propositionnelle ne pourra pas à elle seule rendre compte des mécanismes de raisonnement des mathématiciens.

En fait, même les raisonnements courants impliquent des propositions paramétriques comme le montre l'exemple classique suivant :

– Tous les hommes sont mortels.
– Or, Socrate est un homme.
– Donc, Socrate est mortel.

On voit que la troisième proposition est conséquence logique des deux premières, mais une analyse purement propositionnelle ne rendra pas compte de ce fait. Nous avons donc besoin d'un langage formel plus riche que le calcul des propositions, qui permettra d'exprimer des propriétés vraies pour certains individus pris dans un ensemble, c'est-à-dire des relations.

En mathématique, on définit une *relation* $\mathcal{R}$ d'arité n sur les ensembles $D_1, D_2, \ldots, D_n$ comme un sous-ensemble du produit cartésien $D_1 \times D_2 \times \cdots \times D_n$. Voici à titre d'exemple la description de quelques relations importantes de l'arithmétique :

$$\begin{aligned}\mathcal{PPQ}(x,y) &= \{(x,y) \in (\mathbb{N}\times\mathbb{N}) : x < y\}\\ &= \{(0,1),(0,2),(0,3),\ldots,(1,2),(1,3),\ldots,(2,3),\ldots\}\end{aligned}$$

$$\mathcal{CARRE}(x,y) = \{(x,y) \in (\mathbb{N}\times\mathbb{N}) : y = x^2\} = \{(0,0),(1,1),(2,4),(3,9),\ldots\}$$

$$\mathcal{PR}(x) = \{x \in \mathbb{N} : x \text{ est un nombre premier}\} = \{2,3,5,7,11,13,\ldots\}$$

Définitions. Soit D un ensemble. $\mathcal{R}$ est une *relation* d'arité n *sur le domaine* D si $\mathcal{R}$ est une relation sur D^n. Le *prédicat* R associé à $\mathcal{R}$ est défini par

$$R(d_1, \ldots, d_n) = \mathbf{V} \text{ si et seulement si } (d_1, \ldots, d_n) \in \mathcal{R}.$$

On aura donc

$$PPQ(0,1) = \mathbf{V}\,,\; PPQ(8,4) = \mathbf{F}\,,\; PPQ(3,6) = \mathbf{V}\,,\ldots$$

$$CARRE(0,0) = \mathbf{V}\,,\; CARRE(0,2) = \mathbf{F}\,,\; CARRE(2,4) = \mathbf{V}\;\ldots$$

$$PR(3) = \mathbf{V}\,,\; PR(8) = \mathbf{F}\;\ldots$$

On voit qu'un prédicat est une proposition paramétrique, vraie pour certains éléments d'un domaine et fausse pour les autres.

Il faut souligner d'emblée que le principal apport du calcul des prédicats ne sera pas l'étude des formules paramétriques pour elles-mêmes, mais plutôt l'étude de formules non paramétriques dont certaines composantes sont paramétriques. Pour prendre un exemple célèbre, la question de Fermat n'est pas de savoir si la formule

$$x^n + y^n = z^n \,\wedge\, x,y,z \neq 0 \,\wedge\, n > 2 \tag{1}$$

est vraie pour des entiers x, y, z, n donnés, mais bien de savoir si, oui ou non, il existe un quadruplet d'entiers tel que la formule soit vraie. La première question, du ressort du calcul élémentaire, est clairement paramétrique ; le fait que $2^3+3^3 = 35 \neq 64 = 4^3$ établit clairement que la formule est fausse pour le quadruplet $(2,3,4,3)$ mais ne détermine pas la valeur de vérité pour le quadruplet $(12,13,14,15)$ par exemple. En revanche, le fait que la formule

$$\exists x,y,z,n \in \mathbb{Z}\,[x^n + y^n = z^n \,\wedge\, x,y,z \neq 0 \,\wedge\, n > 2] \tag{2}$$

soit fausse (ce fait a été — avec beaucoup de difficulté — démontré récemment) établit bien que la formule paramétrique précédente est fausse pour tous les quadruplets, et notamment pour $(12,13,14,15)$.

Remarque. Insistons sur le fait que la formule 1 est paramétrique (et sans grand intérêt) alors que la formule 2 ne l'est pas ; il s'agit d'une "honnête" proposition, qui ne peut être que vraie ou

fausse, indépendamment de tout contexte.[70] Cela n'a pas empêché les mathématiciens d'étudier cette formule pendant plus de trois siècles.

Notre introduction à la logique des prédicats se limitera à l'essentiel. On verra d'abord que l'interprétation d'une formule prédicative, quantifiée ou non, implique un domaine de référence D et l'association, à chaque prédicat, d'une relation sur ce domaine. Les constantes individuelles et les variables libres s'interprètent en des éléments de D. On peut aussi introduire des constantes fonctionnelles, dont l'interprétation sera naturellement une fonction qui, à tout n-uplet d'éléments de D, associe un élément de D.

On étudiera ensuite comment les procédures de décision introduites pour la logique propositionnelle s'adaptent à la logique prédicative ; nous verrons que ces techniques (tableaux sémantiques de Beth et Hintikka, séquents de Gentzen, systèmes axiomatiques de Hilbert et résolution de Davis, Putnam et Robinson) donnent lieu à des "semi-procédures" de décision.

5.2 Syntaxe du calcul des prédicats simplifié

Dans un premier temps, nous introduisons les prédicats, les variables et les constantes individuelles, mais pas les fonctions.

5.2.1 Lexique, termes et formules

Soit

- $\mathcal{P} = \{p, q, r, \ldots\}$, un ensemble de symboles arbitraires appelés *symboles de prédicats* (chacun ayant une arité).
  NB : Les propositions atomiques sont des symboles de prédicats d'arité 0.
- $\mathcal{A} = \{a, a_1, a_2, \ldots, b, c, \ldots\}$, un ensemble de symboles arbitraires appelés *constantes* (ou *constantes individuelles*).
- $\mathcal{X} = \{x, x_1, x_2, x', \ldots, y, z, \ldots\}$, un ensemble de symboles arbitraires appelés *variables* (ou *variables individuelles*).

Un *terme* est une constante $a \in \mathcal{A}$ ou une variable $x \in \mathcal{X}$. Une *formule atomique* (ou un *atome*) est une expression $p(t_1, \ldots, t_n)$, où $p \in \mathcal{P}$ est un symbole prédicatif d'arité n et $t_1, \ldots, t_n$ sont des termes. Le concept de *formule* est défini récursivement comme suit.

- Une formule atomique est une formule.
- *true*, *false* sont des formules.
- Si A_1 et A_2 sont des formules, alors $\neg A_1$, $(A_1 \vee A_2)$, $(A_1 \wedge A_2)$, $(A_1 \Rightarrow A_2)$ et $(A_1 \equiv A_2)$ sont des formules.
- Si A est une formule et x une variable, alors $\forall x A$ et $\exists x A$ sont des formules.

Comme dans le cadre propositionnel, on peut justifier l'omission de certaines parenthèses par des règles de précédence. Dans l'ordre décroissant, on a la négation et les quantificateurs, puis la conjonction, la disjonction, l'implication et enfin l'équivalence. Par exemple, la formule $\forall x((\neg\exists y p(x, y)) \vee (\neg\exists y p(y, x)))$ peut se récrire plus simplement en $\forall x(\neg\exists y p(x, y) \vee \neg\exists y p(y, x))$. Néanmoins, l'excès de concision peut nuire à la clarté. Dans la suite, nous utiliserons seulement le fait que la négation et les quantificateurs ont une précédence plus forte

[70] A condition d'accorder aux symboles qui composent la formule leur signification mathématique habituelle.

que les connecteurs binaires. Notons aussi que, dans le cas d'une formule dont l'opérateur principal est un connecteur binaire, il est d'usage d'omettre les parenthèses extérieures.[71]

5.2.2 Portée des quantificateurs, variable libre, variable liée

- La *portée* d'une quantification (d'un quantificateur, d'une variable quantifiée) est la formule à laquelle la quantification s'applique. Dans $\forall x A$ ou dans $\exists x A$, la portée de x (de $\forall x$) est A.[72]
- L'occurrence de la variable x dans la quantification $\forall x$ ou $\exists x$ est dite *quantifiée*.
- Toute occurrence de x dans la portée d'une quantification est dite *liée*.
- Une variable est *libre* si elle n'est ni quantifiée, ni liée.
- Les portées de deux variables x et y sont disjointes ou l'une est incluse dans l'autre.

Ces notions existent aussi en programmation. Considérons l'exemple suivant :

```
program Principal;
var x : integer;

procedure p;
var x : integer;
begin  x := 1;  writeln(x + x)  end;

procedure q;
var y : integer;
begin  y := 1;  writeln(x + y)  end;

begin  x := 5;  p;  q  end.
```

On remarque que les portées de `x` local et de `y` sont disjointes et incluses dans la portée de `x` global. Dans la procédure `q`, on se réfère au `x` global. De même, dans

```
(+ x ((lambda (x) (+ x y)) x))
```

les première et dernière occurrences de `x` sont libres, de même que la variable `y`, tandis que les deuxième et troisième occurrences de `x` sont liées. (On pourrait dire, plus justement, que la deuxième occurrence est "liante" et que la troisième est liée.)

Ces notions apparaissent également en mathématique, et notamment en algèbre et en analyse. Dans l'expression

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} ,$$

[71] Selon la syntaxe adoptée ici, les formules quantifiées et les négations ne comportent pas de paire de parenthèses extérieures.

[72] Les parenthèses extérieures d'une formule dont le connecteur principal est binaire peuvent être omises, mais il n'en découle pas, pour $A =_{def} p(x) \vee q(x)$, que la portée de $\forall x$ dans $\forall x\, p(x) \vee q(x)$ soit $p(x) \vee q(x)$; cette portée est $p(x)$. La formule $\forall x\, A$ doit s'écrire $\forall x\, (p(x) \vee q(x))$.

les variables i et j sont libres, la variable k est liée.[73] Dans l'expression

$$y(x) = y(x_0) + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t\,,$$

la variable x est libre, la variable t est liée.

La distinction entre variable libre et variable liée se fait par simple inspection de la formule considérée. En revanche, la distinction entre constante et variable libre est moins immédiate. Le critère est qu'une variable libre est susceptible d'être quantifiée (et de devenir liée), tandis qu'une constante n'est jamais quantifiable. La distinction se fait par le contexte, ou par des conventions plus ou moins contraignantes et plus ou moins explicites. Dans la formule

$$S = \pi R^2\,,$$

il est "naturel" de considérer π comme la constante bien connue $3.14\ldots$, parce que l'égalité évoque la relation existant entre la surface d'un cercle et son rayon. En revanche, l'égalité

$$V = hb^2$$

évoque la relation entre le volume d'un parallélipipède à base carrée et ses dimensions b et h ; il sera alors tout aussi naturel de considérer h comme une variable libre. La confusion provient des libertés de notation que se permettent les mathématiciens. Les deux formules ci-dessus peuvent se récrire

$$\forall C \in \mathcal{C}\,[S(C) = \pi(R(C))^2]\,,$$

et

$$\forall P \in \mathcal{P}\,[V(P) = h(P)(b(P))^2]\,,$$

ce qui évite toute ambiguïté. Notons cependant que, parfois, le mathématicien est moins laxiste que le logicien. En analyse, on évitera d'écrire

$$y(x) = y(x_0) + \int_{x_0}^{x} f(x, y(x))\,\mathrm{d}x\,,$$

alors qu'en logique il n'est pas interdit d'écrire

$$P(x) \,\wedge\, \forall x\, Q(x)\,,$$

même si nous préférerons

$$P(x) \,\wedge\, \forall u\, Q(u)\,.$$

Considérons quelques exemples.

1. $\varphi_1 =_{def} \forall x \Big(p(x,a) \Rightarrow \exists x\, q(x) \Big)$.
   On préférera éviter d'imbriquer plusieurs quantifications sur la même variable, sans toutefois l'interdire. En fait, on verra que la sémantique de φ_1 est exactement celle de $\forall x \Big(p(x,a) \Rightarrow \exists y\, q(y) \Big)$ ou encore de $\forall y \Big(p(y,a) \Rightarrow \exists x\, q(x) \Big)$.

[73] Selon le contexte, A, B, C et n sont des constantes ou des variables libres.

2. $\varphi_2 =_{def} \exists x \forall x A$.
   Ici aussi, deux quantifications sur x sont imbriquées. La sémantique de φ_2 est celle de $\exists y \forall x A$, pour n'importe quelle variable y sans occurrence (libre) dans A ; la quantification sur y étant inutile, la formule équivaut à $\forall x A$.
3. $\varphi_3 =_{def} \forall x\, p(x,a) \Rightarrow \exists x\, q(x)$.
   Deux variables liées ont le même nom, mais les portées sont disjointes ; il n'y a donc pas de problème.
4. $\varphi_4 =_{def} \forall x\, p(x,a) \Rightarrow q(x)$.
   Une variable libre et une variable liée ont le même nom x. C'est acceptable, mais il est préférable de *renommer* la variable liée et d'écrire, par exemple, $\forall y\, p(y,a) \Rightarrow q(x)$.

5.2.3 Fermetures universelle et existentielle

Une formule est *fermée* ou *close* si elle ne contient aucune variable libre. Lorsqu'une formule A contient les variables libres $x_1, x_2, \ldots, x_n$, on la notera aussi $A(x_1, x_2, \ldots, x_n)$.

Si $x_1, x_2, \ldots, x_n$ sont toutes les variables libres d'une formule A,
– $\forall x_1 \forall x_2 \cdots \forall x_n A$ est la *fermeture universelle* de A.
– $\exists x_1 \exists x_2 \cdots \exists x_n A$ est la *fermeture existentielle* de A.

Exemple. La fermeture universelle de la formule $p(x) \Rightarrow q(x)$ est $\forall x\,(p(x) \Rightarrow q(x))$ et non $\forall x\, p(x) \Rightarrow q(x)$.

5.3 Sémantique du calcul des prédicats

5.3.1 Interprétations

Une *interprétation* $\mathcal{I}$ est un triplet (D, I_c, I_v) tel que :
– D est un ensemble non vide, appelé *domaine d'interprétation* ;
– I_c est une fonction qui associe
  – à toute *constante* a, un objet $I_c[a]$ appartenant à D,
  – à tout symbole prédicatif p (arité n), une relation (arité n) sur D, c'est-à-dire une fonction de D^n dans $\{\mathbf{V}, \mathbf{F}\}$;
– I_v est une fonction qui associe à toute variable x un élément $I_v[x]$ de D.

Voici quatre exemples d'interprétations pour la formule $\forall x\, p(a, x)$:
– $\mathcal{I}_1 = (\mathbb{N}, \quad I_{1c}[p] = \leq, \quad I_{1c}[a] = 0)$;
– $\mathcal{I}_2 = (\mathbb{N}, \quad I_{2c}[p] = \leq, \quad I_{2c}[a] = 1)$;
– $\mathcal{I}_3 = (\mathbb{Z}, \quad I_{3c}[p] = \leq, \quad I_{3c}[a] = 0)$;
– $\mathcal{I}_4 = (\mathcal{S}, \quad I_{4c}[p] = \sqsubseteq, \quad I_{4c}[a] = \lambda)$,
  où $\mathcal{S}$ est l'ensemble des mots sur un alphabet donné ; $w_1 \sqsubseteq w_2$ signifie que w_1 est un préfixe de w_2 ; λ représente le mot vide.

5.3.2 Règles d'interprétation

Des règles sémantiques, appelées *règles d'interprétation*, permettent d'étendre une interprétation à l'ensemble des formules. En fait, une interprétation $\mathcal{I} = (D, I_c, I_v)$ associe

une valeur de vérité à toute formule A et associe un élément de D à tout terme t. En ce qui concerne les termes, on a

– Si x est une variable libre, $\mathcal{I}[x] = I_v[x]$.
– Si a est une constante, $\mathcal{I}[a] = I_c[a]$.

En ce qui concerne les formules, on a

– Si p est un symbole prédicatif d'arité n et si $t_1, \ldots, t_n$ sont des termes, alors $\mathcal{I}[p(t_1, \ldots, t_n)] = (I_c[p])(\mathcal{I}[t_1], \ldots, \mathcal{I}[t_n])$.
– $\mathcal{I}[true] = \mathbf{V}$ et $\mathcal{I}[false] = \mathbf{F}$.
– Si A est une formule, alors $\neg A$ s'interprète comme dans le calcul des propositions, c'est-à-dire $\mathcal{I}[\neg A] = \mathbf{V}$ si $\mathcal{I}[A] = \mathbf{F}$ et $\mathcal{I}[\neg A] = \mathbf{F}$ si $\mathcal{I}[A] = \mathbf{V}$.
– Si A_1 et A_2 sont des formules, alors $(A_1 \vee A_2)$, $(A_1 \wedge A_2)$, $(A_1 \Rightarrow A_2)$, $(A_1 \equiv A_2)$ s'interprètent comme dans le calcul des propositions.
  $\mathcal{I}[(A_1 \wedge A_2)]$ vaut $\mathbf{V}$ si $\mathcal{I}[A_1] = \mathbf{V}$ et $\mathcal{I}[A_2] = \mathbf{V}$, et vaut $\mathbf{F}$ sinon.
  $\mathcal{I}[(A_1 \vee A_2)]$ vaut $\mathbf{V}$ si $\mathcal{I}[A_1] = \mathbf{V}$ ou $\mathcal{I}[A_2] = \mathbf{V}$, et vaut $\mathbf{F}$ sinon.
  $\mathcal{I}[(A_1 \Rightarrow A_2)]$ vaut $\mathbf{V}$ si $\mathcal{I}[A_1] = \mathbf{F}$ ou $\mathcal{I}[A_2] = \mathbf{V}$, et vaut $\mathbf{F}$ sinon.
  $\mathcal{I}[(A_1 \equiv A_2)]$ vaut $\mathbf{V}$ si $\mathcal{I}[A_1] = \mathcal{I}[A_2]$, et vaut $\mathbf{F}$ sinon.

Notation. Si $\mathcal{I} = (D_{\mathcal{I}}, I_c, I_v)$ est une interprétation, si x est une variable et si d est un élément de $D_{\mathcal{I}}$, alors $\mathcal{I}_{x/d}$ désigne l'interprétation $\mathcal{J} = (D_{\mathcal{J}}, J_c, J_v)$ telle que $D_{\mathcal{J}} = D_{\mathcal{I}}$, $J_c = I_c$, $J_v[x] = d$ et $J_v[y] = I_v[y]$ pour toute variable y distincte de x.

– Si A est une formule et x une variable, $\mathcal{I}[\forall x A]$ vaut $\mathbf{V}$ si $\mathcal{I}_{x/d}[A] = \mathbf{V}$ pour tout élément d de D, et vaut $\mathbf{F}$ sinon.
– Si A est une formule et x une variable, $\mathcal{I}[\exists x A]$ vaut $\mathbf{V}$ si $\mathcal{I}_{x/d}[A] = \mathbf{V}$ pour au moins un élément d de D, et vaut $\mathbf{F}$ sinon.

5.3.3 Capture de variable

Les règles d'interprétation des quantifications sont conformes à l'intuition traduite par les noms des quantificateurs. Il faut quand même souligner deux points importants, que l'emploi d'un même lexique pour les variables libres et les variables liées rend délicats :

– La valeur de $\mathcal{I}[\forall x\, A(x)]$ ne dépend pas de $\mathcal{I}[x]$.
– Si $\mathcal{I}[\forall x\, A(x)] = \mathbf{V}$, alors $\mathcal{I}[A(t)] = \mathbf{V}$, pour tout terme t ne donnant lieu à *aucune capture de variable*.
  Exemple. Si $\forall x\, \exists y\, p(x, y)$ est vrai, alors les *instances* $\exists y\, p(a, y)$, $\exists y\, p(x, y)$ et $\exists y\, p(z, y)$ sont nécessairement vraies, mais l'instance $\exists y\, p(y, y)$ peut être fausse. Dans cette instance, l'occurrence y premier argument de p a été capturée et est devenue liée.
  Conclusion. On ne peut pas dire que $\exists y\, p(y, y)$ est une instance licite de $\forall x\, \exists y\, p(x, y)$; on ne peut pas substituer y à x dans $\exists y\, p(x, y)$. Si on veut quand même effectuer cette substitution ou instantiation, on commencera par renommer la variable liée pour éviter la *capture*. On pourra dire, par exemple, que $\exists z\, p(y, z)$ est une instance (après renommage) de $\forall x\, \exists y\, p(x, y)$, ou le résultat de la substitution (après renommage) de y à x dans $\exists y\, p(x, y)$.

On omettra souvent de rappeler que les instantiations et substitutions donnant lieu à capture sont interdites … tout en signalant une fois pour toutes que le phénomène de capture est à la source de nombreuses erreurs !

5.3.4 Satisfaction, modèle

Une formule A *est vraie pour une interprétation* $\mathcal{I}$ ou A *est satisfaite par une interprétation* $\mathcal{I}$ ou $\mathcal{I}$ *est un modèle de* A si $\mathcal{I}[A] = \mathbf{V}$. Cela se note $\models_{\mathcal{I}} A$.

Remarque. On rencontre parfois l'écriture $\mathcal{I} \models A$, mais nous ne l'emploierons pas dans ce cours, pour éviter tout risque de confusion avec l'écriture $U \models A$, introduite au paragraphe suivant.

Exemples. Soit A la formule $\forall x\, p(a, x)$. Les quatre interprétations introduites plus haut attribuent à A une valeur de vérité :

– $D_{\mathcal{I}_1} = \mathbb{N},\ I_{1c}[p] = \leq,\ I_{1c}[a] = 0$; on a $\models_{\mathcal{I}_1} A$.
– $D_{\mathcal{I}_2} = \mathbb{N},\ I_{2c}[p] = \leq,\ I_{2c}[a] = 1$; on a $\not\models_{\mathcal{I}_2} A$.
– $D_{\mathcal{I}_3} = \mathbb{Z},\ I_{3c}[p] = \leq,\ I_{3c}[a] = 0$; on a $\not\models_{\mathcal{I}_3} A$.
– $D_{\mathcal{I}_4} = \mathcal{S},\ I_{4c}[p] = \sqsubseteq,\ I_{4c}[a] = \lambda$; on a $\models_{\mathcal{I}_4} A$.

Définitions. Soit A une formule du calcul des prédicats.

– A est *satisfaisable* ou *consistante* si A a au moins un modèle.
– A est *valide* (cela se note $\models A$) si $\mathcal{I}[A] = \mathbf{V}$ pour toute interprétation $\mathcal{I}$.
– A est *insatisfaisable* ou *inconsistant* si A n'est pas satisfaisable, donc si $\mathcal{I}[A] = \mathbf{F}$ pour toute interprétation $\mathcal{I}$.
– A est *simplement consistante* ou *contingente* si A est consistante mais non valide.

Théorème (dualité validité – consistance). La formule A est valide si et seulement si $\neg A$ est inconsistante.

Exemples.

– $\forall x\, p(a, x)$ est consistante mais non valide.
  $D_{\mathcal{I}_1} = \mathbb{N},\ I_{1c}[p] =\leq,\ I_{1c}[a] = 0 : \models_{\mathcal{I}_1} A$.
  $D_{\mathcal{I}_3} = \mathbb{Z},\ I_{3c}[p] =\leq,\ I_{3c}[a] = 0 : \not\models_{\mathcal{I}_3} A$.
– $\forall x\, p(x) \Rightarrow p(a)$ est valide.
– $\exists x\, p(x) \Rightarrow p(a)$ est simplement consistante.

Remarque. Tout schéma propositionnel valide est aussi un schéma prédicatif valide. Par exemple, du schéma propositionnel valide $\neg\neg A \equiv A$, on peut déduire $\neg\neg(p \wedge q) \equiv (p \wedge q)$, mais aussi $\neg\neg\forall x p(x) \equiv \forall x p(x)$.

5.3.5 Quelques formules valides importantes

– $(\forall x A \wedge \forall x B) \equiv \forall x(A \wedge B)$
– $(\forall x A \vee \forall x B) \Rightarrow \forall x(A \vee B)$
– $\forall x(A \Rightarrow B) \Rightarrow (\forall x A \Rightarrow \forall x B)$
– $\forall x(A \equiv B) \Rightarrow (\forall x A \equiv \forall x B)$

– $\exists x(A \vee B) \equiv (\exists x A \vee \exists x B)$
– $\exists x(A \wedge B) \Rightarrow (\exists x A \wedge \exists x B)$
– $\exists x(A \Rightarrow B) \equiv (\forall x A \Rightarrow \exists x B)$

– $\forall x A \equiv \neg(\exists x \neg A)$
– $\forall x A \Rightarrow \exists x A$

– $\forall x \forall y A \equiv \forall y \forall x A$

– $\exists x \exists y A \equiv \exists y \exists x A$
– $\exists x \forall y A \Rightarrow \forall y \exists x A$

On observera que le remplacement d'une implication par une équivalence produit, dans chaque cas, une formule non valide. Considérons par exemple le cas de la formule $\exists x(A \wedge B) \Rightarrow (\exists x A \wedge \exists x B)$. Il est évident, vu la règle sémantique se rapportant à l'existentielle, que, si $C \Rightarrow D$ est valide, alors $\exists x C \Rightarrow \exists x D$ est valide. En conséquence, les deux formules $\exists x(A \wedge B) \Rightarrow \exists x A$ et $\exists x(A \wedge B) \Rightarrow \exists x B$ sont valides. D'autre part, si $C \Rightarrow D$ et $C \Rightarrow E$ sont vraies ou valides, alors $C \Rightarrow (D \wedge E)$ est vraie ou valide. Il en découle que $\exists x(A \wedge B) \Rightarrow (\exists x A \wedge \exists x B)$ est valide.

Pour montrer que l'implication inverse (ou réciproque, ou converse) n'est pas valide, il suffit d'en donner un antimodèle. On prend pour domaine l'ensemble $\mathbb{N}$; $A(x)$ est interprété en "x est pair" et $B(x)$ en "x est impair". La formule $\exists x A \wedge \exists x B$ est vraie : elle signifie qu'il existe au moins un entier naturel pair, et au moins un entier naturel impair. La formule $\exists x(A \wedge B)$ est fausse : elle signifie qu'il existerait au moins un entier naturel à la fois pair et impair.

Notons enfin que le passage des quantifications informelles aux quantifications formelles (et réciproquement) est un exercice important, souvent facile, mais parfois délicat. Considérons un exemple :

Toutes les licornes sont dangereuses, donc il existe une licorne dangereuse.

Une modélisation hâtive telle que

$$\forall x LD(x) \Rightarrow \exists x LD(x)$$

pourrait laisser croire à la validité du raisonnement informel, ce qui serait incorrect. En effet, les licornes n'existent pas ; on peut donc les qualifier sans erreur de dangereuses (ou d'inoffensives), mais on ne peut pas affirmer qu'il existe une licorne, dangereuse ou non. Le paradoxe apparent disparaît si l'on utilise un modèle formel correct, à savoir

$$\forall x[L(x) \Rightarrow D(x)] \Rightarrow \exists x[L(x) \wedge D(x)]$$

Cette dernière formule est consistante mais n'est pas valide.

5.3.6 Conséquence logique, équivalence logique

Définitions. Soit U un ensemble de formules et soient A et B deux formules.

– A est une *conséquence logique* de U (cela se note $U \models A$) si A est vrai dans tous les modèles de U.
  Remarque. En pratique, l'ensemble U sera souvent un ensemble de formules fermées. On a alors $U \models A$ si et seulement si $U \models \forall x\, A$.
– A et B sont *logiquement équivalentes* (cela se note $A \leftrightarrow B$) si $\mathcal{I}[A] = \mathcal{I}[B]$ pour toutes les interprétations $\mathcal{I}$.

Comme dans le calcul des propositions, on a $\models A$ si et seulement si $\emptyset \models A$.

Théorème. Une formule est valide si et seulement si sa fermeture universelle est valide ; une formule est consistante si et seulement si sa fermeture existentielle est consistante.

Théorème. Deux formules A et B sont logiquement équivalentes si et seulement si la formule $A \equiv B$ est valide.

Remarque. Ces théorèmes découlent immédiatement des définitions et des règles d'interprétation. On notera que, si x est la seule variable libre de $A(x)$, alors $A(x)$ et $A(y)$ ne

sont en général pas logiquement équivalentes, mais $\forall x\, A(x)$ et $\forall y\, A(y)$ le sont. On évite des complications sans perdre d'expressivité en considérant les problèmes de validité, de consistance et de conséquence logique seulement pour les formules fermées. Lors de la fermeture d'une formule, l'ordre des quantifications n'a pas d'importance (c'est pourquoi on parle de "la" fermeture universelle ou existentielle d'une formule).

Théorème de l'échange. Soit A une sous-formule d'une formule B et soit A' une formule telle que $A \leftrightarrow A'$. Soit B' la formule résultant du remplacement de A par A' dans B. On a $B \leftrightarrow B'$.
Démonstration. Comme dans le cas propositionnel, on procède par induction structurelle. Les seuls cas inductifs nouveaux sont liés à la quantification. Pour la quantification universelle, on doit seulement montrer que si $B(x) \leftrightarrow B'(x)$, on a aussi $\forall x\, B(x) \leftrightarrow \forall x\, B'(x)$, ce qui est évident.

5.4 Le théorème de compacité

Le théorème de compacité subsiste en logique prédicative, et un ensemble de formules est consistant si et seulement si tous ses sous-ensembles finis sont consistants. On peut prouver ce résultat important en adaptant la preuve donnée dans le cadre propositionnel, mais nous verrons un moyen plus rapide plus loin.

6 Analyse des formules prédicatives

6.1 Méthode simple pour formules simples

En logique propositionnelle, l'application directe des règles sémantiques permet toujours d'analyser une formule, c'est-à-dire de déterminer si elle est valide, contingente ou inconsistante. La méthode des tables de vérité concrétise cette approche fondamentalement simple. En logique prédicative, la situation est moins favorable parce qu'une formule consistante admet souvent une infinité de modèles très différents. Néanmoins, si on accepte certaines restrictions sur l'emploi des quantificateurs, l'approche sémantique directe reste possible.

6.1.1 Formules sans quantification

Une formule sans quantification est une combinaison booléenne de formules atomiques. De telles formules peuvent s'analyser par la méthode des tables de vérité, si on assimile tout atome à une proposition élémentaire. Considérons par exemple la formule Φ :

$$P(a,a) \wedge \neg P(a,x) \wedge Q(a,b) \wedge (Q(a,a) \Rightarrow P(a,x))\,.$$

La formule Φ comporte quatre atomes syntaxiquement distincts qui, par ordre d'occurrence, sont $P(a,a)$, $P(a,x)$, $Q(a,b)$ et $Q(a,a)$. La *version propositionnelle* de Φ s'obtient en substituant uniformément à ces quatre atomes les propositions élémentaires distinctes, par exemple p_1, p_2, p_3 et p_4, respectivement, ce qui donne

$$p_1 \wedge \neg p_2 \wedge p_3 \wedge (p_4 \Rightarrow p_2)$$

On note immédiatement les lemmes suivants :
Lemme 1. Toute formule sans quantification admet une version propositionnelle unique.
Lemme 2. Si Φ est une formule sans quantification, la version propositionnelle de $\neg\Phi$ est la négation de la version propositionnelle de Φ.
Remarque. Deux formules sans quantification distinctes peuvent avoir la même version propositionnelle.
Définition. Une formule sans quantification est *p-valide* (resp. p-consistante, p-contingente) si sa version propositionnelle est valide (resp. consistante, contingente).
Exemple. La formule Φ donnée plus haut est p-contingente, puisque sa version propositionnelle est contingente.[74]
Lemme 3. Une formule sans quantification est valide (resp. consistante, contingente) si et seulement si elle est p-valide (resp. p-consistante, p-contingente).
Démonstration. Vu le lemme 2, il suffit de démontrer qu'une formule sans quantification Φ admet un modèle si et seulement si sa version propositionnelle admet un modèle.
La condition est nécessaire. Soit I un modèle de Φ. L'interprétation I attribue une valeur de vérité à chacun des atomes de Φ. Soit J l'interprétation telle que $J(p_k)$ est la valeur associée par I au $k$ième atome de Φ ; l'interprétation J est un modèle de la version propositionnelle de Φ.

[74]Cette version propositionnelle admet en fait un modèle et quinze anti-modèles.

La condition est suffisante. Soit J un modèle de la version propositionnelle de Φ. On construit un modèle I de Φ comme suit. Le domaine d'interprétation se compose des constantes et des variables de Φ. La fonction d'interprétation I applique chaque terme sur lui-même. Il reste à définir $I(P)$, pour tout prédicat P intervenant dans Φ. Si P est, par exemple, d'arité 2, il faut définir, vu le choix que nous avons fait pour D, $I(P(d_1,d_2))$ pour tous $d_1,d_2 \in D$. On distingue deux cas : si $P(d_1,d_2)$ est le $k$ième atome de Φ, on pose $I(P(d_1,d_2)) = J(p_k)$, sinon on choisit (arbitrairement) $I(P(d_1,d_2)) = \mathbf{V}$.

6.2 Méthode des tableaux sémantiques

Dans le cadre prédicatif comme dans le cadre propositionnel, la méthode des tableaux sémantiques consiste en une recherche systématique des modèles. Pour déterminer si la formule A est valide, on recherche un modèle de $\neg A$; si un tel modèle n'existe pas, A est valide ; s'il en existe (au moins) un, A n'est pas valide. De plus, la méthode des tableaux sémantiques est analytique : elle réduit une formule à ses composants. En ce sens, les composants d'une formule universelle $\forall x\, A(x)$ seront les formules $A(c)$, où c est n'importe quel terme ; le composant d'une formule universelle $\exists x\, A(x)$ sera la formule $A(a)$, où a est une constante inédite, appelée paramètre. Le traitement de la quantification étant délicat, nous en illustrerons d'abord les dangers. On se limite à l'étude des formules fermées, ce qui n'est pas une réelle restriction.

6.2.1 Quelques exemples

Exemple 1 (naïf). Test de validité de $\forall x \Big(p(x) \Rightarrow q(x)\Big) \Rightarrow \Big(\forall x\, p(x) \Rightarrow \forall x\, q(x)\Big)$.

Dans le tableau 47, on a d'abord instancié $\neg\forall x q(x)$ (formule existentielle, équivalente à $\exists x\, \neg q(x)$), en $\neg q(a)$. On a ensuite instancié les formules universelles $\forall x\, p(x)$ et $\forall x\, (p(x) \Rightarrow q(x))$ en $p(a)$ et $p(a) \Rightarrow q(a)$. Intuitivement, une existentielle sera instanciée une seule fois (par branche), en utilisant une constante spécifique ; au contraire, une universelle pourra être instanciée plusieurs fois, au moyen de toutes les constantes disponibles. Pour rendre ceci rigoureux, il faudra préciser les mots "pourra", "spécifique" et "disponible".

Exemple 2 (incorrect !). Test de validité de $\forall x \Big(p(x) \vee q(x)\Big) \Rightarrow \Big(\forall x\, p(x) \vee \forall x\, q(x)\Big)$.

Le tableau 48 montre simplement que sa racine n'admet pas de modèle à un élément. En revanche, elle admet un modèle à deux éléments (ce que le tableau ne montre pas ; il est donc incorrect !) et la formule testée n'est donc pas valide. Le problème est lié au choix de la même constante a dans l'instantiation de $\neg\forall x\, p(x)$ et de $\neg\forall x\, q(x)$. Cette identité est abusive : le fait que les formules $p(x)$ et $q(x)$ admettent chacune des "contre-exemples" n'impliquent pas qu'elles admettent des contre-exemples communs.

Exemple 2 (version corrigée). Test de validité de $\forall x\, (p(x) \vee q(x)) \Rightarrow (\forall x\, p(x) \vee \forall x\, q(x))$.
On recommence donc la dérivation, en utilisant pour les existentielles deux constantes distinctes a et b. Cela implique naturellement que l'universelle soit instanciée au moyen de a *et* de b. Le tableau de la figure 49 comporte une branche ouverte à laquelle correspond un modèle $\mathcal{I}$ de la racine, tel que $\mathcal{I}[p(a)] = \mathcal{I}[q(b)] = \mathbf{V}$ et $\mathcal{I}[p(b)] = \mathcal{I}[q(a)] = \mathbf{F}$. Cette interprétation montre que la formule testée n'est pas valide (cf. fig. 49).

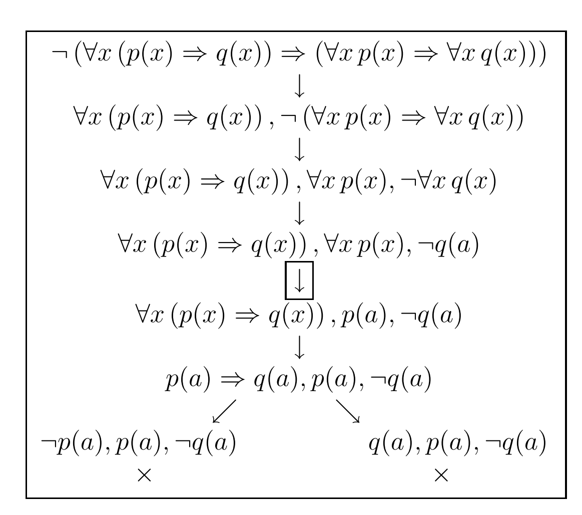


FIG. 47 – Exemple 1 (naïf) ; l'étape encadrée est fautive.

$$\neg(\forall x\,(p(x) \vee q(x)) \Rightarrow (\forall x\, p(x) \vee \forall x\, q(x)))$$
$$\downarrow$$
$$\forall x\,(p(x) \vee q(x))\,, \neg(\forall x\, p(x) \vee \forall x\, q(x))$$
$$\downarrow$$
$$\forall x\,(p(x) \vee q(x))\,, \neg\forall x\, p(x), \neg\forall x\, q(x)$$
$$\downarrow$$
$$\forall x\,(p(x) \vee q(x))\,, \neg\forall x\, p(x), \neg q(a)$$
$$\boxed{\downarrow}$$
$$\forall x\,(p(x) \vee q(x))\,, \neg p(a), \neg q(a)$$
$$\downarrow$$
$$p(a) \vee q(a), \neg p(a), \neg q(a)$$
$$\swarrow \qquad \searrow$$
$$p(a), \neg p(a), \neg q(a) \qquad q(a), \neg p(a), \neg q(a)$$
$$\times \qquad\qquad \times$$

FIG. 48 – Exemple 2, tableau incorrect ; l'étape encadrée est fautive.

Exemple 3. Test de $\forall x \exists y\, p(x,y) \ \wedge\ \forall x \neg p(x,x) \wedge \forall x \forall y \forall z\, (p(x,y) \wedge p(y,z) \Rightarrow p(x,z))$.
Le tableau sémantique de la figure 50 est infini. Son unique branche doit être considérée comme ouverte car elle définit un modèle (nécessairement infini) de la formule testée.

Exemple 4. Test de validité pour la formule
$\forall x \exists y\, p(x,y) \ \wedge\ \forall x \neg p(x,x) \wedge \forall x \forall y \forall z\, (p(x,y) \wedge p(y,z) \Rightarrow p(x,z)) \ \wedge\ \forall x\, (q(x) \wedge \neg q(x))$.
Si, dans le tableau 51, on instanciait indéfiniment $\forall x \exists y\, p(x,y)$, en négligeant à tort les autres formules, la branche ne se fermerait pas et serait infinie.

Les exemples 1 et 2 suggèrent que l'instantiation des existentielles, ou *exemplification*,

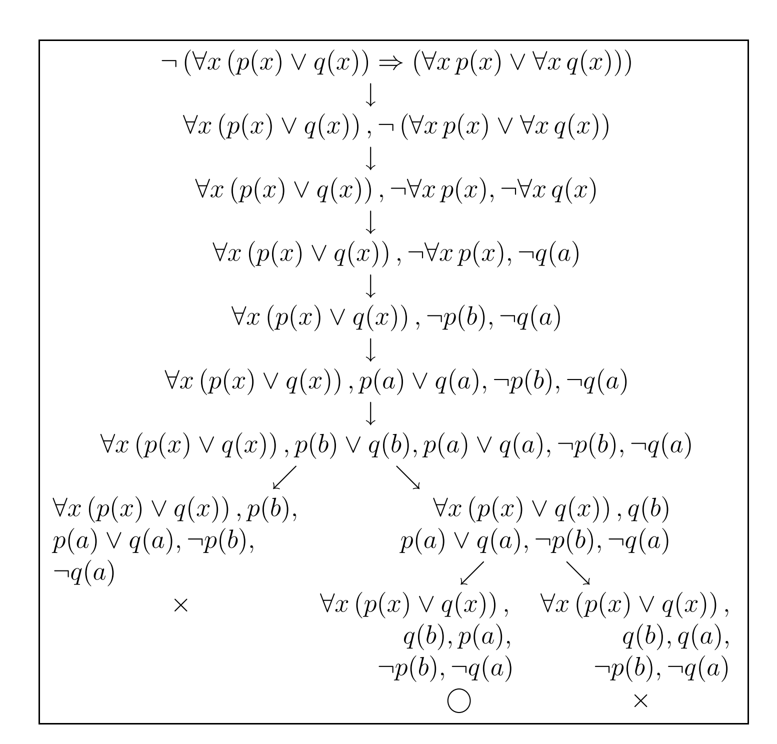


FIG. 49 – Exemple 2, tableau correct.

se fasse au moyen de constantes *inédites*, appelées aussi *paramètres*. Cela n'exclut pas les formules à "petits" modèles : en l'absence du prédicat spécial d'égalité, si $\{a, b\}$ par exemple est le domaine d'un modèle de A, $\{a, a_1, \ldots, b, b_1, \ldots\}$ donnera aussi lieu à un modèle, si les a_i et b_j sont des "clones" de a et b, c'est-à-dire tels que φ, $\varphi[a/a_i]$ et $\varphi[b/b_j]$ aient même valeur de vérité, pour toute formule φ.

L'exemple 3 montre que la construction d'un tableau sémantique peut ne pas se terminer, en particulier si la formule étudiée est consistante mais n'admet que des modèles infinis. On espère néanmoins que la méthode permettra toujours de reconnaître les formules inconsistantes, négations de formules valides.

L'exemple 4 indique enfin que cette inconsistance pourrait n'être pas reconnue si les règles de décomposition n'étaient pas appliquées de manière "équitable" ; il faut notamment se méfier de la règle *générative* d'instantiation des universelles, qui peut s'appliquer indéfiniment. On *doit* l'appliquer à toute constante introduite par la règle d'exemplification (sauf si la branche se ferme).

6.2.2 Règles de décomposition

Aux règles propositionnelles α et β s'ajoutent les règles prédicatives γ et δ.

– *Règles de prolongation (type α) et de ramification (type β)*

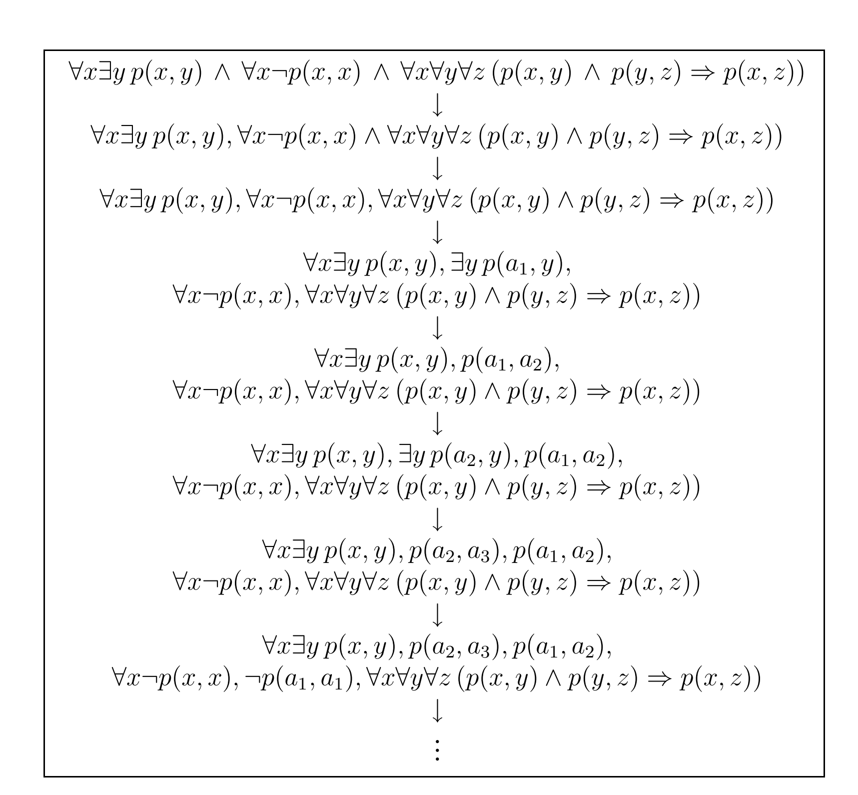


FIG. 50 – Exemple 3, tableau infini.

α	α_1	α_2
$A_1 \wedge A_2$	A_1	A_2
$\neg(A_1 \vee A_2)$	$\neg A_1$	$\neg A_2$
$\neg(A_1 \Rightarrow A_2)$	A_1	$\neg A_2$
$\neg(A_1 \Leftarrow A_2)$	$\neg A_1$	A_2

β	β_1	β_2
$B_1 \vee B_2$	B_1	B_2
$\neg(B_1 \wedge B_2)$	$\neg B_1$	$\neg B_2$
$B_1 \Rightarrow B_2$	$\neg B_1$	B_2
$B_1 \Leftarrow B_2$	B_1	$\neg B_2$

– *Règles génératives (type γ) et exemplatives (type δ)*

γ	$\gamma(c)$
$\forall x\, A(x)$	$A(c)$
$\neg\exists x\, A(x)$	$\neg A(c)$

(constante c quelconque)
(constante a inédite)

δ	$\delta(a)$
$\exists x\, A(x)$	$A(a)$
$\neg\forall x\, A(x)$	$\neg A(a)$

Rappelons aussi la règle d'élimination des doubles négations.

6.2.3 Construction d'un tableau sémantique

On présente d'abord l'algorithme, qui est non déterministe, puis des restrictions à ce non-déterminisme ; ces restrictions sont nécessaires pour assurer la terminaison (dans certains cas) et la complétude.

Algorithme de construction.

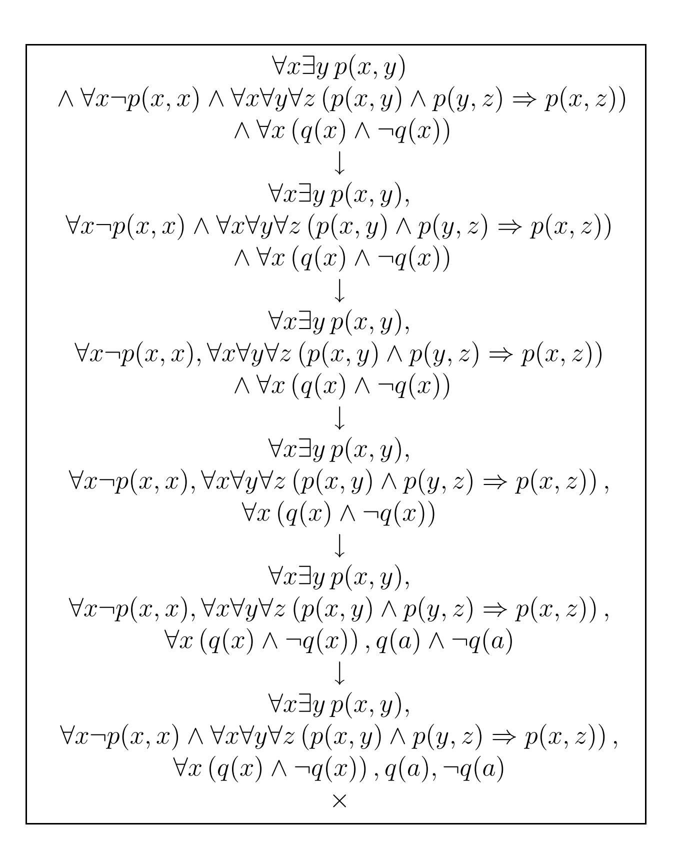


FIG. 51 – Exemple 4.

Initialisation : une racine étiquetée $\{A\}$.

Etape inductive : sélectionner une feuille non marquée ℓ ; soit $U(\ell)$ son étiquette.

- Si $U(\ell)$ contient une paire complémentaire, alors marquer ℓ comme *fermée* '$\times$' ;
- Si $U(\ell)$ ne contient que des littéraux (sans paire complémentaire), alors marquer ℓ comme *ouverte* '$\bigcirc$' ;
- Si $U(\ell)$ n'est pas un ensemble de littéraux, sélectionner une formule dans $U(\ell)$:
  - si c'est une α-formule A, créer un nouveau nœud ℓ', descendant de ℓ, et étiqueter ℓ' avec $U(\ell') = (U(\ell) \setminus \{A\}) \cup \{\alpha_1, \alpha_2\}$;
  - si c'est une β-formule B, créer deux nouveaux nœuds ℓ' et ℓ'', descendants de ℓ, et étiqueter ℓ' avec $U(\ell') = (U(\ell) \setminus \{B\}) \cup \{\beta_1\}$ et étiqueter ℓ'' avec $U(\ell'') = (U(\ell) \setminus \{B\}) \cup \{\beta_2\}$;
  - si c'est une γ-formule C, créer un nouveau nœud ℓ', descendant de ℓ, et étiqueter ℓ' avec $U(\ell') = U(\ell) \cup \{\gamma(c)\}$;
  - si c'est une δ-formule D, créer un nouveau nœud ℓ', descendant de ℓ, et étiqueter ℓ' avec $U(\ell') = (U(\ell) \setminus \{D\}) \cup \{\delta(a)\}$, où a est une constante *qui n'apparaît pas* dans

$U(\ell)$.[75]
Terminaison : survient quand toutes les feuilles sont marquées.

Règles additionnelles de construction.

Le non-déterminisme de l'algorithme de construction intervient

1. lors du choix du nœud à développer ;
2. lors du choix de la formule à décomposer dans ce nœud ;
3. lors du choix du terme c lors d'une γ-réduction.[76]

Il faut adopter une stratégie qui garantisse les deux conditions suivantes.

– Toute formule qui apparaît sur une branche ouverte de l'arbre se voit appliquer une règle de décomposition quelque part sur cette branche.
  Autrement dit, toute formule décomposable est décomposée, à moins que la branche se ferme.
– Pour toute γ-formule A et toute constante a qui apparaissent sur une branche ouverte, une règle d'instantiation est appliquée à la formule A avec la constante a quelque part sur cette branche.
  Toute constante apparaissant sur une branche est utilisée à un moment donné pour instancier les γ-formules sur cette branche, à moins qu'elle se ferme.

Un moyen simple et classique d'assurer le respect des conditions d'équité est d'étiqueter les nœuds par des *listes* de formules. Le(s) nœud(s) successeurs de n est (sont) obtenus par "décomposition" de la première formule de la liste $U(n)$ non réduite à un littéral ; la liste $U(n')$ (et $U(n'')$, s'il y a lieu) est obtenue en supprimant de $U(n)$ la formule traitée, et en ajoutant en fin de liste le(s) "composant(s)" adéquats. Dans le cas d'une formule générative, la formule supprimée en tête de liste est réinsérée en queue de liste.

Une autre méthode appropriée est la suivante. Lorsqu'une règle générative est activée, on construit immédiatement les instances correspondant à toutes les constantes introduites jusque là dans la branche. De même, quand une exemplification est faite, ce qui provoque l'adjonction dans la branche d'une constante inédite, on "réactive" les γ-réductions déjà accomplies, pour insérer les instances correspondant à cette nouvelle constante. Ceci nécessite une gestion organisée de l'ensemble des constantes et des activations de règles génératives.

La stratégie n'est pas nécessaire pour obtenir l'adéquation, mais elle l'est pour obtenir la complétude. En effet, la stratégie ne vise qu'à éviter le report définitif de réductions susceptibles de fermer une branche. Le point est d'ailleurs délicat, puisque la construction d'un tableau sémantique peut ne pas se terminer.

[75]On voit que cette constante n'apparaît pas non plus dans l'étiquette d'un ancêtre de ℓ ; cette contrainte devrait être introduite explicitement si on convenait de ne pas récrire les littéraux étiquetant un nœud dans l'étiquette de ses successeurs (convention que l'on adopte parfois pour alléger la construction). Dans ce cas, il convient de préciser que la recherche de paires complémentaires se fait dans toute la branche, et non seulement dans son dernier nœud.

[76]Lors d'une δ-réduction, la constante choisie doit être inédite ; on a vu que le non-respect de cette condition rendait la méthode inadéquate (exemple 1). En revanche, le choix du nom de cette constante inédite est clairement sans importance ; les δ-réductions, au contraire des γ-réductions, n'introduisent donc pas de vrai non-déterminisme.

Rappelons enfin que certaines règles de priorité permettent souvent d'accélérer la construction du tableau. En particulier, on effectuera les α-réductions avant les β-réductions, pour limiter le nombre de branchements. On évitera d'instancier une γ-formule par une constante inédite (c'est inutile) sauf naturellement dans le cas où aucune δ-réduction n'a pu être effectuée. L'exemple 5 (fig. 52) illustre certaines de ces règles. Il illustre aussi un point délicat. La formule $\forall x \exists y\, r(x,y) \Rightarrow \exists y \forall x\, r(x,y)$, où r est un prédicat binaire, est non valide ; on en déduit naturellement que, si $R(x,y)$ est une formule quelconque admettant x et y comme variables libres, la formule $\forall x \exists y\, R(x,y) \Rightarrow \exists y \forall x\, R(x,y)$ est *généralement* non valide. Pour certains choix de R, la formule peut cependant être valide ; c'est le cas notamment si $R(x,y)$ est $p(x) \Rightarrow q(y)$.

$$
\begin{array}{c}
\neg\Big(\forall x \exists y\,(p(x) \Rightarrow q(y)) \Rightarrow \exists y \forall x\,(p(x) \Rightarrow q(y))\Big) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\forall x\,(p(x) \Rightarrow q(c))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg(p(a) \Rightarrow q(c))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ p(a)\,,\ \neg q(c)\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\exists y\,(p(a) \Rightarrow q(y))\,,\ p(a)\,,\ \neg q(c) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
p(a) \Rightarrow q(b)\,,\ p(a)\,,\ \neg q(c) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\forall x\,(p(x) \Rightarrow q(b))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
p(a) \Rightarrow q(b)\,,\ p(a)\,,\ \neg q(c) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
\neg(p(d) \Rightarrow q(b))\,,\ p(a) \Rightarrow q(b)\,,\ p(a)\,,\ \neg q(c) \\
\downarrow \\
\forall x \exists y\,(p(x) \Rightarrow q(y))\,,\ \neg\exists y \forall x\,(p(x) \Rightarrow q(y)) \\
p(d)\,,\ \neg q(b)\,,\ p(a) \Rightarrow q(b)\,,\ p(a)\,,\ \neg q(c) \\
\swarrow \qquad\qquad \searrow \\
\begin{array}{cc}
\forall x \exists y\,(\ldots)\,,\ \neg\exists y \forall x\,(\ldots) & \forall x \exists y\,(\ldots)\,,\ \neg\exists y \forall x\,(\ldots) \\
p(d),\ \neg q(b),\ \neg p(a),\ p(a),\ \neg q(c) & p(d),\ \neg q(b),\ q(b),\ p(a),\ \neg q(c) \\
\downarrow & \downarrow \\
\times & \times
\end{array}
\end{array}
$$

FIG. 52 – Exemple 5.

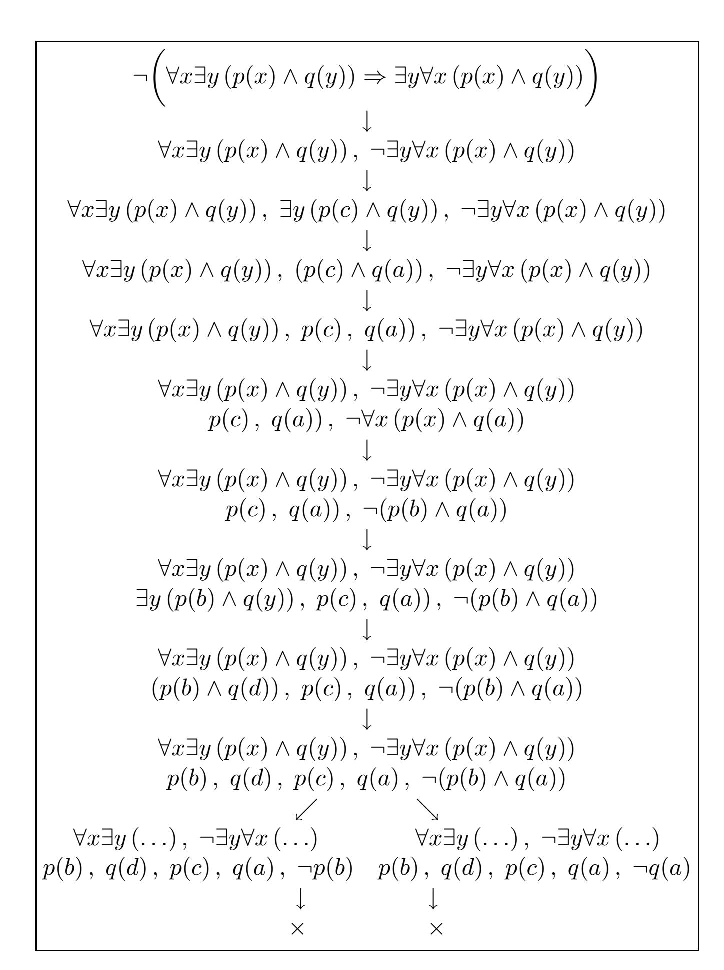


FIG. 53 – Exemple 6.

6.2.4 Adéquation de la méthode des tableaux sémantiques

Théorème. Soit $T(A)$ un tableau sémantique dont la racine est A. Si $T(A)$ est fermé,[77] alors la formule A est inconsistante.

Remarque. Vu que tout sous-arbre d'un tableau fermé est aussi un tableau fermé, on prouvera un résultat apparemment plus fort, à savoir que les étiquettes de tous les nœuds (pas seulement la racine) d'un tableau fermé sont inconsistantes.

Démonstration. On prouve par induction sur la hauteur h du nœud n dans $T(A)$ que l'étiquette de n est un ensemble inconsistant. Ce sera vrai en particulier pour la racine de l'arbre, dont l'étiquette est le singleton $\{A\}$.

[77] c'est-à-dire si toutes les branches de $T(A)$ sont fermées.

– Cas de base, $h = 0$: le nœud n est une feuille, nécessairement fermée, donc $U(n)$ contient une paire complémentaire et est inconsistant.
– Cas inductif, $h > 0$: une règle α, β, γ ou δ a été utilisée pour créer le(s) descendant(s) du nœud n. Les cas α et β sont les mêmes que dans la démonstration de la version propositionnelle du théorème. On considère successivement les cas γ et δ.
  – **Règle** γ : n : $\{\forall x\, A(x)\} \cup U_0$ $\downarrow$ n' : $\{\forall x\, A(x), A(c)\} \cup U_0$
    $U(n')$ est inconsistant par hypothèse inductive, donc $U(n)$ est inconsistant ; en effet, tout modèle de $U(n)$ serait aussi un modèle de $U(n')$, ou s'étendrait immédiatement en un tel modèle, au cas où la constante c n'interviendrait pas dans U_0.
  – **Règle** δ : n : $\{\exists x\, A(x)\} \cup U_0$ $\downarrow$ n' : $\{A(a)\} \cup U_0$
    où a est une constante qui n'apparaît pas dans $U(n)$. Si $U(n)$ était consistant, il existerait une interprétation $\mathcal{I} = (D, I_c, I_v)$ telle que $\mathcal{I}[\exists x\, A(x)] = \mathbf{V}$, donc il existerait $d \in D$ tel que $\mathcal{I}_{x/d}[A(x)] = \mathbf{V}$.
    Définissons $\mathcal{J} = (D, J_c, I_v)$ avec J_c obtenu en étendant[78] I_c de sorte que $J_c[a] = d$. Alors, $\mathcal{J}[A(a)] = \mathbf{V}$ et $\mathcal{J}[U_0] = \mathcal{I}[U_0] = \mathbf{V}$, donc $\mathcal{J}$ satisfait $U(n')$, une contradiction.

6.2.5 Complétude de la méthode des tableaux sémantiques

On abordera la complétude comme dans le cas propositionnel, via la notion d'ensemble de Hintikka.

Ensembles de Hintikka. *Définition.* Soit U un ensemble de formules fermées, et C_U l'ensemble des constantes individuelles ayant au moins une occurrence dans U. L'ensemble U est un *ensemble de Hintikka* si les cinq conditions suivantes sont satisfaites :

1. Si A est une formule atomique, on a $A \notin U$ ou $\neg A \notin U$.
2. Si $\alpha \in U$ est une α-formule, alors $\alpha_1 \in U$ et $\alpha_2 \in U$.
3. Si $\beta \in U$ est une β-formule, alors $\beta_1 \in U$ ou $\beta_2 \in U$.
4. Si γ est une γ-formule, alors pour tout $a \in C_U$ on a $\gamma(a) \in U$.
5. Si δ est une δ-formule, alors il existe $a \in C_U$ tel que $\delta(a) \in U$.

Théorème. Soit *b* une branche ouverte d'un tableau T construit en respectant les conditions d'équité. L'ensemble $U = \bigcup_{n\in b} U(n)$ est de Hintikka.

Démonstration. La condition d'ouverture assure le respect par U de la condition 1. Les règles α, β, γ et δ *permettent* l'insertion dans U des éléments requis par les conditions 2, 3, 4 et 5, respectivement. La stratégie de construction *impose* que tout élément ajoutable soit effectivement ajouté.

[78] On est sûr de pouvoir procéder à l'extension puisque a est une constante inédite, telle que $I_c[a]$ n'existe pas.

Remarque. La branche *b* peut être infinie ; dans ce cas, elle est nécessairement ouverte et elle définit un modèle infini.

Lemme de Hintikka. Tout ensemble de Hintikka est consistant.

Démonstration. Soit U un ensemble de Hintikka. Le *modèle canonique* $\mathcal{I}_U = (D, I_c, I_v)$ associé à U est défini comme suit :

- $D = \{a, b, \ldots, \}$ est l'ensemble des constantes apparaissant dans les formules de U ;
- On construit la fonction d'interprétation I_c comme suit :
  - Pour toute constante $d \in D$, on pose $I_c[d] = d$.
  - Pour tout symbole prédicatif p (arité m) apparaissant dans U, on pose
    $I_c[p](I_c[a_1], \ldots, I_c[a_m]) = \mathbf{V}$, si $p(a_1, \ldots, a_m) \in U$,
    $I_c[p](I_c[a_1], \ldots, I_c[a_m]) = \mathbf{F}$, si $\neg p(a_1, \ldots, a_m) \in U$.
    $I_c[p](I_c[a_1], \ldots, I_c[a_m])$ est arbitraire si $\{p(a_1, ..., a_m), \neg p(a_1, ..., a_m)\} \cap U = \emptyset$.
- I_v est quelconque, puisqu'il n'y a pas de variables libres.

Il reste à montrer que pour toute formule (fermée) $A \in U$, on a $\mathcal{I}[A] = \mathbf{V}$. Cela se fait par induction sur la structure de A. (Exercice.)

Complétude. Comme dans le cas propositionnel, on peut montrer que si un tableau sémantique (respectant la stratégie de construction) est ouvert, alors la formule étiquetant la racine est consistante. Un modèle est le modèle canonique de Hintikka associé à une branche ouverte. On en déduit que si A est une formule inconsistante, tout tableau $T(A)$ (respectant la stratégie de construction) est fermé. Rappelons que, dans le cadre prédicatif, une branche peut être infinie. Cependant, si on respecte la stratégie de construction, une branche infinie est nécessairement ouverte.

Pour analyser une formule (fermée) A, on peut construire les tableaux sémantiques $T(A)$ et $T(\neg A)$. Si $T(A)$ est fermé (et donc fini), A est inconsistant. Si $T(\neg A)$ est fermé (et donc fini), A est valide. Si $T(A)$ et $T(\neg A)$ sont ouverts, A et $\neg A$ sont simplement consistants.

Dans le cas propositionnel, l'analyse se termine toujours. Dans le cas prédicatif, l'analyse *peut* ne pas se terminer si A et $\neg A$ sont simplement consistants. Cela n'a rien d'étonnant ; contrairement au calcul des propositions, le calcul des prédicats n'est que semi-décidable.

6.3 Méthode des séquents

6.3.1 Dualité entre séquents et tableaux

Comme dans le cas propositionnel, on peut "par dualité" obtenir une dérivation de séquent au départ d'un tableau sémantique.

Un exemple suffira à rappeler le procédé. La validité de la formule

$$(\forall x\, p(x) \vee \forall x\, q(x)) \Rightarrow \forall x\, (p(x) \vee q(x))$$

est démontrée par la méthode des tableaux (figure 54) puis par celle des séquents sans antécédent (figure 55). D'une figure à l'autre, l'arbre est retourné et chaque formule est remplacée par son complément. Les feuilles fermées deviennent des séquents valides ou *axiomes* ; les feuilles ouvertes deviennent des séquents non valides ou *hypothèses*. Dans les tableaux sémantiques, les ensembles de formules sont conjonctifs et la virgule a donc valeur

conjonctive. Dans les séquents, la virgule a valeur disjonctive quand elle se trouve à droite de la flèche, dans le succédent. Un séquent peut aussi avoir un antécédent, dans lequel la virgule a valeur conjonctive. On ne change pas la sémantique d'un séquent en faisant passer l'une de ses formules du succédent vers l'antécédent ou réciproquement, à condition de changer sa polarité. Par exemple, les quatre séquents ci-dessous sont équivalents :

$$\rightarrow A, \neg B \qquad B \rightarrow A \qquad \neg A \rightarrow \neg B \qquad \neg A, B \rightarrow$$

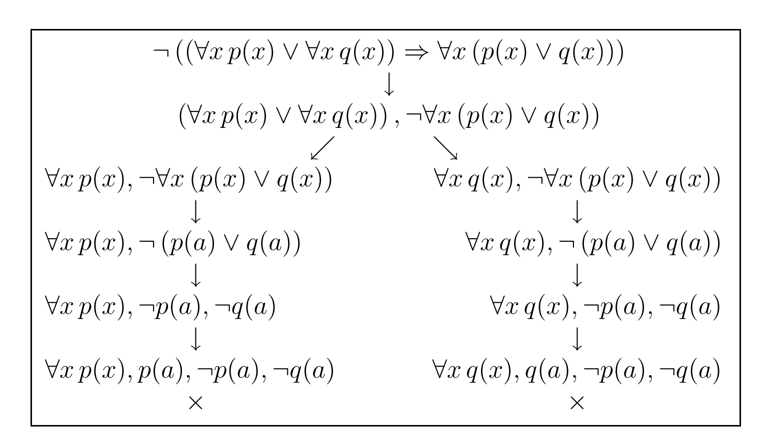


FIG. 54 – Un tableau sémantique ...

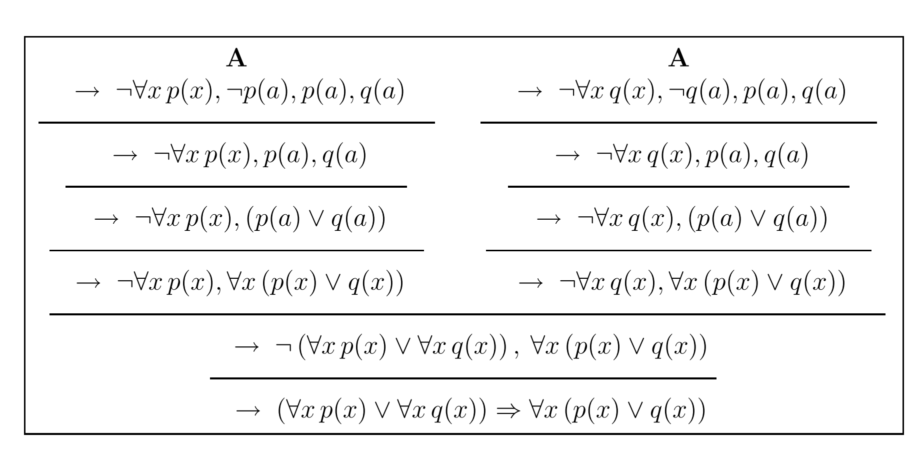


FIG. 55 – ... et la dérivation de séquent duale.

6.3.2 Règles du système de Gentzen

Comme dans le cas propositionnel, un séquent est un axiome si le même atome (variante : la même formule) apparaît dans l'antécédent et dans le succédent. En outre, les règles α et β

$$\dfrac{\dfrac{\dfrac{\dfrac{\mathbf{A}}{\forall x\, p(x), p(a) \;\rightarrow\; p(a), q(a)}}{\forall x\, p(x) \;\rightarrow\; p(a), q(a)}}{\forall x\, p(x) \;\rightarrow\; (p(a) \vee q(a))}}{\forall x\, p(x) \;\rightarrow\; \forall x\, (p(x) \vee q(x))} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\mathbf{A}}{\forall x\, q(x), q(a) \;\rightarrow\; p(a), q(a)}}{\forall x\, q(x) \;\rightarrow\; p(a), q(a)}}{\forall x\, q(x) \;\rightarrow\; (p(a) \vee q(a))}}{\forall x\, q(x) \;\rightarrow\; \forall x\, (p(x) \vee q(x))}$$

$$\dfrac{(\forall x\, p(x) \vee \forall x\, q(x)) \;\rightarrow\; \forall x\, (p(x) \vee q(x))}{\rightarrow\; (\forall x\, p(x) \vee \forall x\, q(x)) \Rightarrow \forall x\, (p(x) \vee q(x))}$$

FIG. 56 – Dérivation, séquents avec antécédent.

sont les mêmes que dans le cadre propositionnel. Nous rappelons seulement celles relatives à l'implication.

– *règle* α :
$$\frac{U, A \;\rightarrow\; V, B}{U \;\rightarrow\; V,\, (A \Rightarrow B)}$$
– *règle* β :
$$\frac{U \;\rightarrow\; V, A \qquad U, B \;\rightarrow\; V}{U,\, (A \Rightarrow B) \;\rightarrow\; V}$$

On ajoute des règles génératives (règles γ) et les règles d'exemplification (règles δ), pour traiter les formules quantifiées :

– *règles* γ :
$$\exists : \frac{U \;\rightarrow\; V,\, \exists x\, A(x),\, A(c)}{U \;\rightarrow\; V,\, \exists x\, A(x)}$$
$$\forall : \frac{U,\, \forall x\, A(x), A(c) \;\rightarrow\; V}{U,\, \forall x\, A(x) \;\rightarrow\; V}$$
– *règles* δ :
$$\forall : \frac{U \;\rightarrow\; V,\, A(a)}{U \;\rightarrow\; V,\, \forall x\, A(x)} \quad \text{si } a \text{ n'apparaît pas dans la conclusion.}$$
$$\exists : \frac{U,\, A(a) \;\rightarrow\; V}{U,\, \exists x\, A(x) \;\rightarrow\; V} \quad \text{si } a \text{ n'apparaît pas dans la conclusion.}$$

La dérivation de la figure 55 (séquents sans antécédent) est reprise dans le cadre général à la figure 56. La figure 57 montre la dérivation relative à une formule importante. On voit à la figure 58 comment l'analyse d'une formule non valide peut conduire à une séquence infinie. La figure 59 illustre le danger du non-respect de la restriction attachée à la règle δ. C'est cette restriction qui empêcherait ici la fermeture (incorrecte). La dérivation de la figure 59 montre que la formule est vraie dans un domaine réduit à un élément, sans mettre en évidence le fait

— essentiel — que la même formule est le plus souvent fausse dans un domaine comportant plusieurs éléments.

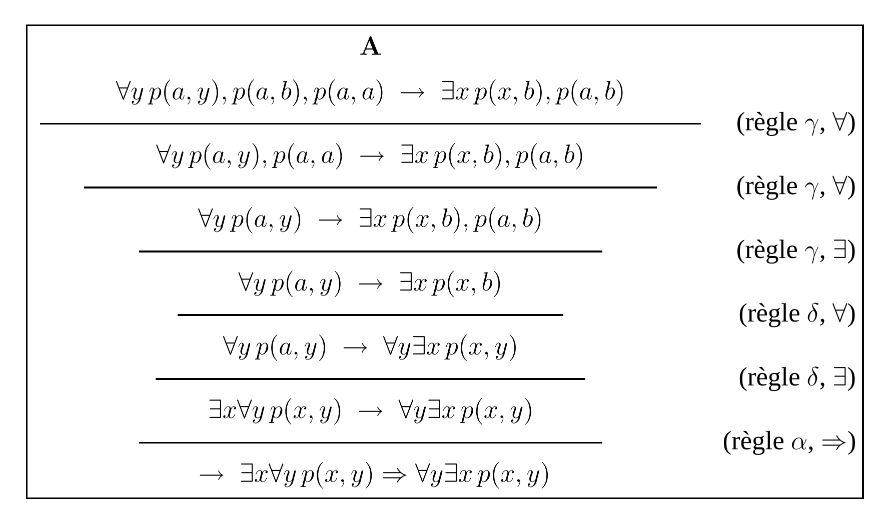


FIG. 57 – Validité de $\exists x \forall y\, p(x, y) \Rightarrow \forall y \exists x\, p(x, y)$.

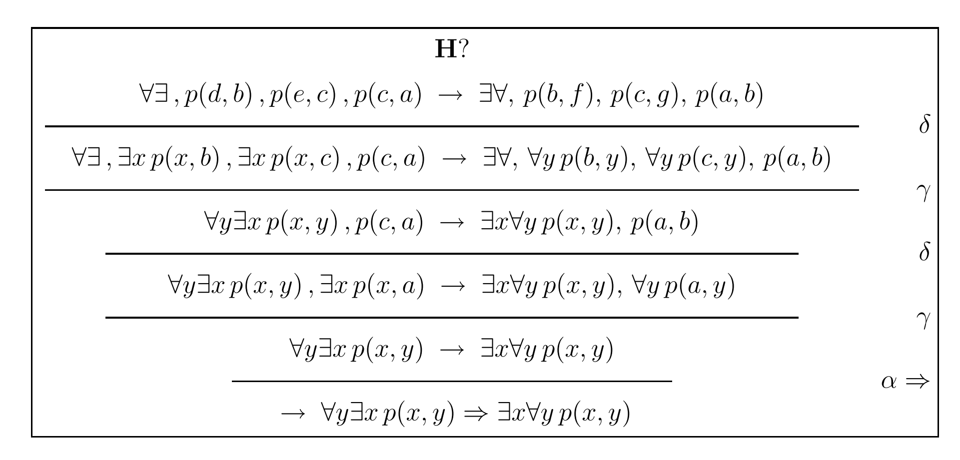


FIG. 58 – Non-validité de $\forall y \exists x\, p(x, y) \Rightarrow \exists x \forall y\, p(x, y)$.

6.3.3 Propriétés du système de Gentzen

Adéquation et complétude. Une formule A est valide si et seulement si elle est racine d'une dérivation de séquent finie dont toutes les feuilles sont des axiomes.

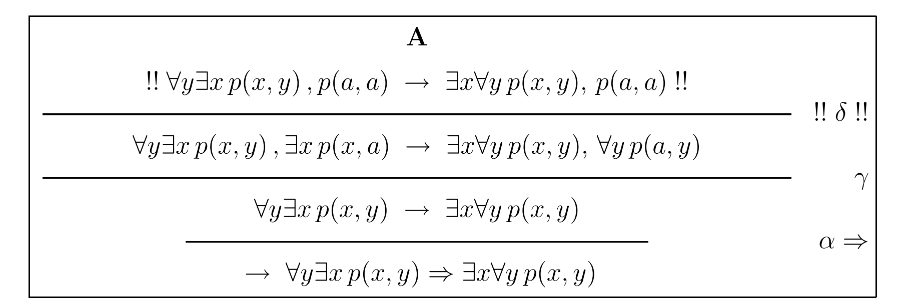


FIG. 59 – Dérivation incorrecte de $\forall y \exists x\, p(x,y) \Rightarrow \exists x \forall y\, p(x,y)$.

Terminaison. L'obtention d'une dérivation adéquate exige le respect d'une stratégie équitable, comme pour les tableaux sémantiques. Malgré cela, l'analyse d'une formule non valide peut donner lieu à une dérivation infinie.
Analyticité. Une règle est *analytique* si tous les composants (formules et sous-formules) des prémisses apparaissent dans la conclusion. Les règles α et β sont analytiques. L'idée sous-jacente est que la découverte de prémisse(s) appropriée(s) au départ de la conclusion doit être triviale. En ce sens, on peut considérer que les règles δ sont analytiques. Pour les règles γ, le choix de la constante c devient critique s'il peut être effectué d'une infinité de manières. Ce sera le cas pour le calcul des prédicats avec symboles fonctionnels (pour l'instant, c ne peut être qu'une constante individuelle).
Réversibilité. Tout modèle des prémisses d'une règle (correcte) est aussi un modèle de sa conclusion. Une règle est *réversible* si la réciproque est également vraie. Les règles α, β et γ sont réversibles. Les règles δ sont "quasi réversibles" : tout modèle de la conclusion peut être *étendu* en un modèle de la prémisse. Dans tous les cas, la validité de la conclusion implique celle de la ou des prémisse(s).
Remarque. La *correction* des règles δ n'est pas évidente ; elle dépend crucialement de la condition imposée à la constante a.

6.4 Digression : le syllogisme catégorique

La méthode des tableaux sémantiques permet d'analyser des formules et des ensembles de formules avec une grande facilité et surtout une grande généralité. La logique prédicative est un domaine où l'approche formelle et symbolique s'est révélée très enrichissante. Un moyen de s'en rendre compte est d'essayer d'analyser des raisonnements prédicatifs sans passer par l'étape de formalisation : on recule très vite devant la difficulté de la tâche. La théorie classique des syllogismes catégoriques est, dans une certaine mesure, une brillante exception. Cette théorie concerne un fragment réduit, mais très important en pratique, de la logique prédicative. A l'intérieur de ce fragment, on dispose de règles simples et efficaces permettant de décider de la validité ou de la consistance de formules et ensembles de formules.

La théorie originale, informelle mais dans l'ensemble rigoureuse et claire,[79] est intéressante

[79] Malgré quelques imprécisions que nous discuterons.

en soi, mais son étude exhaustive sortirait du cadre de ce cours. En revanche, il est utile ici de situer cette théorie dans le cadre général.

6.4.1 La formule de base et ses variantes

Etant donnés deux prédicats unaires[80] P et Q (dans cet ordre), on appelle *formule de base* la formule $\forall x\,(P(x) \Rightarrow Q(x))$. Les *variantes* s'obtiennent en introduisant des négations, portant sur le conséquent du conditionnel ou sur toute la formule. On a donc quatre possibilités, souvent identifiées par les quatre lettres **A**, **E**, **I** et **O** :[81]

$\boxed{\forall x\,(P(x) \Rightarrow Q(x))}$	**A**	universelle affirmative	$\neg\exists x\,(P(x) \wedge \neg Q(x))$
$\boxed{\forall x\,(P(x) \Rightarrow \neg Q(x))}$	**E**	universelle négative	$\neg\exists x\,(P(x) \wedge Q(x))$
$\neg\forall x\,(P(x) \Rightarrow \neg Q(x))$	**I**	particulière affirmative	$\boxed{\exists x\,(P(x) \wedge Q(x))}$
$\neg\forall x\,(P(x) \Rightarrow Q(x))$	**O**	particulière négative	$\boxed{\exists x\,(P(x) \wedge \neg Q(x))}$

Ces quatre formules sont dites "de type PQ". Une formule dont le type est PQ ou QP est une $\{P,Q\}$*-formule* (il y a donc huit $\{P,Q\}$-formules).

6.4.2 Le syllogisme catégorique

Le "jeu du syllogisme catégorique" consiste à choisir une $\{P,Q\}$-formule, une $\{Q,R\}$-formule et une $\{P,R\}$-formule, puis à déterminer si la troisième formule (la *conclusion*) est conséquence logique des deux premières (les *prémisses*). Il y a donc $8^3 = 512$ possibilités. Dans la mesure où les rôles de P et de R sont interchangeables, on peut imposer que la conclusion soit de type PR (et non de type RP), ce qui élimine la moitié des possibilités. On appelle alors *mineure* la $\{P,Q\}$-prémisse, et *majeure* la $\{Q,R\}$-prémisse ; les termes $P(x)$, $Q(x)$ et $R(x)$ sont dits respectivement *mineur*, *moyen* et *majeur*.[82] Le syllogisme catégorique est la règle d'inférence

$$\frac{\textit{majeure} \qquad \textit{mineure}}{\textit{conclusion}}$$

Etant donnés les trois prédicats P, Q et R, il existe donc 256 syllogismes catégoriques. Le problème est de déterminer lesquels sont valides, c'est-à-dire tels que la conclusion soit conséquence logique des prémisses.

[80] ou monadiques, c'est-à-dire d'arité 1.
[81] Pour "**A**ff**I**rmo" et "n**E**g**O**".
[82] La conclusion comporte le mineur *puis* le majeur. La prémisse mineure comporte le mineur et le moyen, la prémisse majeure comporte le majeur et le moyen ; dans les prémisses, l'ordre des deux termes n'est pas imposé. Notons aussi l'emploi classique du mot "terme", dont l'acception moderne est différente ; $P(x)$, $Q(x)$ et $R(x)$ sont en fait des *atomes*, ou *formules atomiques*.

6.4.3 Esquisse de l'approche classique du problème

L'approche classique du problème du syllogisme, développée par Aristote et ses successeurs, visait à synthétiser un petit nombre de règles permettant de déterminer si un syllogisme était valide ou non. Voici cinq exemples de telles règles, que nous acceptons provisoirement comme "intuitivement évidentes" :

1. Si les deux prémisses sont négatives, le syllogisme n'est pas valide.
2. Si les deux prémisses sont particulières, le syllogisme n'est pas valide.
3. Si une prémisse est particulière, la conclusion doit être particulière.
4. Si une prémisse est négative, la conclusion doit être négative.
5. Si les deux prémisses sont affirmatives, la conclusion doit être affirmative.

Une approche systématique consiste à organiser les 256 possibilités selon divers critères. Classiquement, on utilise les notions de *figure* et de *mode*. La figure est fonction du type des prémisses et, plus précisément, de l'ordre des termes dans les prémisses. Il y a donc quatre figures, reprises dans le tableau suivant :

Figure :	première	deuxième	troisième	quatrième
Majeure	QR	RQ	QR	RQ
Mineure	PQ	PQ	QP	QP
Conclusion	PR	PR	PR	PR

Le mode d'un syllogisme est déterminé par la nature des prémisses et de la conclusion. Par exemple, le mode AEI désigne le cas où la *majeure* est universelle affirmative (A), la *mineure* est universelle négative (E) et la *conclusion* est particulière affirmative (I). Il y a donc $4^3 = 64$ modes possibles, chacun pouvant exister dans les quatre figures. Cependant, les cinq règles mentionnées plus haut réduisent fortement ce nombre. Par exemple, la première des règles intuitives permet d'éliminer des modes tels que EEE et OEO ; la deuxième permet d'éliminer III (entre autres) ; la troisième provoque notamment le rejet de AIA ; des modes tels que AOI et AIO contreviennent respectivement aux quatrième et cinquième règles. En fait, ces cinq règles ne laissent subsister que douze modes susceptibles de donner lieu à des syllogismes valides ; ce sont

AAA , AAI , AEE , AEO , AII , AOO ,
EAE , EAO , EIO , IAI , IEO , OAO .

Il ne reste donc que $12 \times 4 = 48$ syllogismes potentiellement valides ; parmi ceux-ci, nous allons voir que 15 syllogismes seulement sont valides, selon l'acception moderne du concept de validité.

6.4.4 Les diagrammes de Venn

C'est par une démarche informelle qu'Aristote et les scolastiques ont déterminé quels syllogismes étaient valides et lesquels ne l'étaient pas. Les syllogismes valides ont reçu des noms conventionnels, dont les voyelles rappellent le mode. Par exemple, le raisonnement classique "Tous les Grecs sont des humains, tous les humains sont mortels, donc tous les Grecs sont mortels" est un syllogisme dont on détecte aisément la structure :[83]

(M) *Tous les humains sont mortels.*
(m) *Tous les Grecs sont des humains.*
(C) *Tous les Grecs sont mortels.*

Ce syllogisme appartient à la première figure et au mode AAA ; cette combinaison se note AAA-1 ou, de manière plus classique (et plus poétique), BARBARA. Les trois "A" rappellent que les prémisses et la conclusion sont toutes trois des universelles affirmatives. De même, le syllogisme

(M) *Tous les étudiants sont intelligents.*
(m) *Certains humains ne sont pas intelligents.*
(C) *Certains humains ne sont pas des étudiants.*

appartient à la deuxième figure ; c'est un exemple de AOO-2 ou BAROCO, la majeure étant universelle affirmative, la mineure et la conclusion étant particulières négatives.

Un moyen simple et concret d'appréhender la validité d'un syllogisme consiste à utiliser un diagramme de Venn à trois composants (figure 60). Le cercle de gauche représente le mineur $P(x)$, celui de droite le majeur $R(x)$, le cercle du haut correspondant au moyen $Q(x)$. Ces cercles déterminent huit zones numérotées de 0 à 7. Tout objet appartient à l'une de ces zones, selon la valeur de vérité qu'il attribue au mineur, au majeur et au moyen. Par exemple, la zone 4 est intérieure aux cercles mineur et moyen, mais extérieure au cercle majeur ; elle regroupe donc les objets rendant vrais le mineur et le moyen, mais faux le majeur.

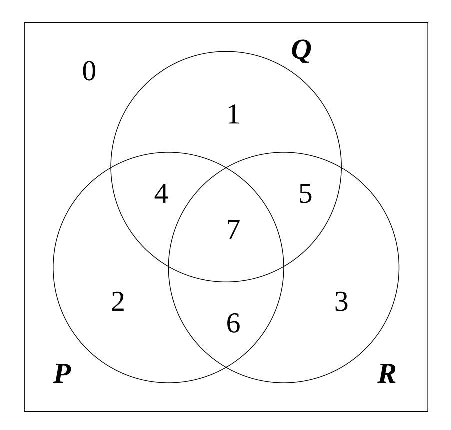


FIG. 60 – Diagramme de Venn

Les prémisses et conclusions des syllogismes correspondent à des assertions de vacuité ou de non-vacuité de certaines zones ; un syllogisme sera valide si son "interprétation graphique" est correcte. Nous illustrons cette technique de vérification par quelques exemples et contre-exemples.

[83]Nous convenons d'énoncer systématiquement la majeure avant la mineure, quoique l'ordre des prémisses soit sans influence sur la validité d'un raisonnement.

Le syllogisme AAA-1 que nous venons d'évoquer s'analyse aisément. La prémisse majeure $\forall x\,(Q(x) \Rightarrow R(x))$ signifie que tout objet vérifiant le moyen vérifie aussi le majeur, donc que les zones 1 et 4 sont vides, ce que nous notons $1 \cup 4 = \emptyset$; de même, la prémisse mineure $\forall x\,(P(x) \Rightarrow Q(x))$ devient $2 \cup 6 = \emptyset$. A la conclusion $\forall x\,(P(x) \Rightarrow R(x))$ correspond l'assertion $2 \cup 4 = \emptyset$. Il est clair que, si les zones 1, 4, 2 et 6 sont vides, alors les zones 2 et 4 sont vides ; on en déduit la validité de BARBARA.

Considérons encore le cas AOO-2. La prémisse majeure $\forall x\,(R(x) \Rightarrow Q(x))$ devient $3 \cup 6 = \emptyset$ et la prémisse mineure $\exists x\,(P(x) \wedge \neg Q(x))$ devient $2 \cup 6 \neq \emptyset$; la conclusion $\exists x\,(P(x) \wedge \neg R(x))$ devient $2 \cup 4 \neq \emptyset$. A nouveau, il est clair que si les zones 3 et 6 sont toutes deux vides et que les zones 2 et 6 ne sont pas toutes deux vides, alors la zone 2 n'est pas vide, et donc les zones 2 et 4 ne sont pas toutes deux vides, ce qui établit la validité de BAROCO.

Le diagramme de Venn permet aussi de voir pourquoi un syllogisme n'est pas valide. Considérons par exemple AIO-4. La prémisse majeure $\forall x\,(R(x) \Rightarrow Q(x))$ devient $3 \cup 6 = \emptyset$ et la prémisse mineure $\exists x\,(Q(x) \wedge P(x))$ devient $4 \cup 7 \neq \emptyset$; on ne peut pas déduire de cela $2 \cup 4 \neq \emptyset$, correspondant à la conclusion $\exists x\,(P(x) \wedge \neg R(x))$. En effet, la situation où 2, 3, 4 et 6 sont vides tandis que 7 ne l'est pas vérifie les deux prémisses mais pas la conclusion.

Les diagrammes de Venn peuvent aussi être utilisés pour l'étude des cinq règles intuitives données plus haut. Considérons par exemple la deuxième règle : "si les deux prémisses sont particulières, le syllogisme n'est pas valide". En effet, une prémisse particulière se traduit par l'assertion "les zones x et y ne sont pas vides toutes les deux". De deux assertions de ce type, on ne peut rien déduire d'intéressant.

6.4.5 Taxonomie des syllogismes catégoriques

Il suffit de passer en revue les 48 syllogismes "potentiellement valides" et d'appliquer la technique du diagramme de Venn à chacun d'eux pour isoler les quinze syllogismes valides. La plupart des auteurs mentionnent cependant plus de quinze syllogismes valides, parce qu'ils acceptent comme valide, au moins dans certains cas, le mécanisme de *subalternation*. La *subalterne* d'une PQ-formule universelle est la PQ-formule particulière (ou existentielle) correspondante.[84] Le mécanisme de subalternation consiste à déduire la subalterne de l'universelle correspondante.

Ce mécanisme n'est pas valide stricto sensu, puisque nous avons déjà signalé au paragraphe 5.3.5 (à propos de l'exemple des licornes) que la subalterne n'était pas conséquence logique de l'universelle ; on a

$$\forall x[L(x) \Rightarrow D(x)] \not\models \exists x[L(x) \wedge D(x)]\,.$$

Cependant, on peut, au moyen d'une prémisse additionnelle, obtenir une version correcte de la subalternation :

$$\{\exists x\, L(x)\,,\ \forall x[L(x) \Rightarrow D(x)]\} \models \exists x[L(x) \wedge D(x)]\,.$$

Dans le langage naturel, on peut parfois considérer que la prémisse manquante est implicite. Nous qualifierons de *quasi-valide* un syllogisme dont la validité dépend de l'emploi de la subalternation, et donc de l'existence d'une prémisse implicite. Cette prémisse sera toujours de la même nature : elle affirme l'existence d'un objet au moins vérifiant le majeur, le moyen ou le mineur. Dans le cadre des diagrammes de Venn, cette prémisse prend donc l'une des trois formes suivantes :

(Moyen) $1 \cup 4 \cup 5 \cup 7 \neq \emptyset$;
(Mineur) $2 \cup 4 \cup 6 \cup 7 \neq \emptyset$;
(Majeur) $3 \cup 5 \cup 6 \cup 7 \neq \emptyset$.

Le tableau récapitulatif des syllogismes valides ou quasi-valides est représenté à la figure 61.[85]

Les quinze syllogismes valides sont AAA-1, EAE-1, AII-1 et EIO-1 pour la première figure, AEE-2, EAE-2, AOO-2 et EIO-2 pour la deuxième figure, AII-3, IAI-3, EIO-3 et OAO-3 pour la troisième figure, et AEE-4, IAI-4 et EIO-4, pour la quatrième figure. Dans le tableau, les noms anciens ont été utilisés ; on obtient la nomenclature moderne en ne retenant que les voyelles.

Cinq syllogismes valides, BARBARA, CELARENT, CAMESTRES, CESARE et CAMENES, ont une conclusion universelle ; en remplaçant celle-ci par sa subalterne, on obtient respectivement les cinq syllogismes quasi-valides BARBARI, CELARO, CAMESTROS, CESARO et CAMENOS. Les dix autres syllogismes valides, dont la conclusion est particulière, ont également une prémisse particulière ;[86] en remplaçant celle-ci par sa superalterne, on obtient aussi des syllogismes quasi-valides, dont quatre distincts des précédents. Tout d'abord, DARII, FERIO, BAROCO et FESTINO redonnent respectivement BARBARI, CELARO, CAMESTROS et CESARO. En outre, dans la troisième figure, DATISI et DISAMIS donnent chacun DARAPTI tandis que BOCARDO et FERISON donnent chacun FELAPTON. Enfin, dans la quatrième figure, DIMARIS donne BRAMANTIP et FRESISON donne FESAPO. On a donc en tout neuf syllogismes quasi-valides.

6.4.6 Théorie ancienne de la réduction, syllogismes valides

Les syllogismes valides de la première figure peuvent sembler "plus naturels" que ceux des autres figures, sans doute parce que les langages naturels, et notamment le français, favorisent l'expression d'un raisonnement syllogistique dans cette figure. Cela explique le point de vue classique, consistant à justifier la validité des syllogismes des autres figures en ramenant ceux-ci à des syllogismes de première figure, au moyen de transformations "évidentes". Ces transformations sont la *conversion simple*, représentée par la lettre S, la *(per)mutation des prémisses*, représentée par la lettre M et la méthode *par contradiction*, représentée par la lettre C. Les trois méthodes sont valides selon la logique moderne et se basent sur des équivalences logiques. Pour la conversion simple, qui s'aplique aux affirmatives particulières et aux négatives universelles, il s'agit de

$$\exists x\,[P(x) \wedge Q(x)] \simeq \exists x\,[Q(x) \wedge P(x)]$$

et de

$$\forall x\,[P(x) \Rightarrow \neg Q(x)] \simeq \forall x\,[Q(x) \Rightarrow \neg P(x)]\,.$$

[84] On dit parfois aussi que la formule universelle est la *superalterne* de la formule particulière.

[85] La dénomination "quasi-valide" n'appartient pas à la terminologie usuelle ; nous l'entendons du raisonnement privé de sa prémisse implicite existentielle. En effet, si nous tenons compte de celle-ci, le raisonnement n'est plus, stricto sensu, un syllogisme ; en revanche, il devient valide.

[86] jamais les deux (règle 2 du paragraphe 6.4.3).

Première figure

BARBARA $\begin{array}{c}\forall x\,[Q(x) \Rightarrow R(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \forall x\,[P(x) \Rightarrow R(x)]\end{array}$	CELARENT $\begin{array}{c}\forall x\,[Q(x) \Rightarrow \neg R(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \forall x\,[P(x) \Rightarrow \neg R(x)]\end{array}$
DARII $\begin{array}{c}\forall x\,[Q(x) \Rightarrow R(x)]\\ \exists x\,[P(x) \wedge Q(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$	FERIO $\begin{array}{c}\forall x\,[Q(x) \Rightarrow \neg R(x)]\\ \exists x\,[P(x) \wedge Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$
BARBARI $\begin{array}{c}\exists x\, P(x)\\ \forall x\,[Q(x) \Rightarrow R(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$	CELARO $\begin{array}{c}\exists x\, P(x)\\ \forall x\,[Q(x) \Rightarrow \neg R(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$

Deuxième figure

CAMESTRES $\begin{array}{c}\forall x\,[R(x) \Rightarrow Q(x)]\\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\\ \hline \forall x\,[P(x) \Rightarrow \neg R(x)]\end{array}$	CESARE $\begin{array}{c}\forall x\,[R(x) \Rightarrow \neg Q(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \forall x\,[P(x) \Rightarrow \neg R(x)]\end{array}$
BAROCO $\begin{array}{c}\forall x\,[R(x) \Rightarrow Q(x)]\\ \exists x\,[P(x) \wedge \neg Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$	FESTINO $\begin{array}{c}\forall x\,[R(x) \Rightarrow \neg Q(x)]\\ \exists x\,[P(x) \wedge Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$
CAMESTROS $\begin{array}{c}\exists x\, P(x)\\ \forall x\,[R(x) \Rightarrow Q(x)]\\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$	CESARO $\begin{array}{c}\exists x\, P(x)\\ \forall x\,[R(x) \Rightarrow \neg Q(x)]\\ \forall x\,[P(x) \Rightarrow Q(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$

Troisième figure

DATISI $\begin{array}{c}\forall x\,[Q(x) \Rightarrow R(x)]\\ \exists x\,[Q(x) \wedge P(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$	DISAMIS $\begin{array}{c}\exists x\,[Q(x) \wedge R(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$
BOCARDO $\begin{array}{c}\exists x\,[Q(x) \wedge \neg R(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$	FERISON $\begin{array}{c}\forall x\,[Q(x) \Rightarrow \neg R(x)]\\ \exists x\,[Q(x) \wedge P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$
DARAPTI $\begin{array}{c}\exists x\, Q(x)\\ \forall x\,[Q(x) \Rightarrow R(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$	FELAPTON $\begin{array}{c}\exists x\, Q(x)\\ \forall x\,[Q(x) \Rightarrow \neg R(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$

Quatrième figure

CAMENES $\begin{array}{c}\forall x\,[R(x) \Rightarrow Q(x)]\\ \forall x\,[Q(x) \Rightarrow \neg P(x)]\\ \hline \forall x\,[P(x) \Rightarrow \neg R(x)]\end{array}$	DIMARIS $\begin{array}{c}\exists x\,[R(x) \wedge Q(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$
CAMENOS $\begin{array}{c}\exists x\, P(x)\\ \forall x\,[R(x) \Rightarrow Q(x)]\\ \forall x\,[Q(x) \Rightarrow \neg P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$	FRESISON $\begin{array}{c}\forall x\,[R(x) \Rightarrow \neg Q(x)]\\ \exists x\,[Q(x) \wedge P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$
BRAMANTIP $\begin{array}{c}\exists x\, R(x)\\ \forall x\,[R(x) \Rightarrow Q(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge R(x)]\end{array}$	FESAPO $\begin{array}{c}\exists x\, Q(x)\\ \forall x\,[R(x) \Rightarrow \neg Q(x)]\\ \forall x\,[Q(x) \Rightarrow P(x)]\\ \hline \exists x\,[P(x) \wedge \neg R(x)]\end{array}$

FIG. 61 – Tableau des syllogismes valides ou quasi-valides

Pour la permutation des prémisses, il s'agit de

$$A, B \models C \text{ si et seulement si } B, A \models C$$

et pour la méthode par contradiction, il s'agit de

$$A, B \models C \text{ si et seulement si } A, \neg C \models \neg B$$

et de

$$A, B \models C \text{ si et seulement si } \neg C, B \models \neg A\,.$$

Les consonnes S, M et C, apparaissant après une voyelle dans le nom d'un syllogisme des autres figures, indiquent les transformations à appliquer pour se ramener au syllogisme de la première figure partageant la même initiale. Par exemple, l'initiale C nous dit que CAMESTRES se ramène à CELARENT ; les deux "S" de CAMESTRES indiquent des conversions simples à opérer sur la mineure et sur la conclusion, ce qui transforme

$$\forall x\,[R(x) \Rightarrow Q(x)]\,,\ \forall x\,[P(x) \Rightarrow \neg Q(x)] \ \models\ \forall x\,[P(x) \Rightarrow \neg R(x)]$$

en

$$\forall x\,[R(x) \Rightarrow Q(x)]\,,\ \forall x\,[Q(x) \Rightarrow \neg P(x)] \ \models\ \forall x\,[R(x) \Rightarrow \neg P(x)]\,;$$

en outre, le "M" de CAMESTRES indique une permutation des prémisses, ce qui conduit à

$$\forall x\,[Q(x) \Rightarrow \neg P(x)]\,,\ \forall x\,[R(x) \Rightarrow Q(x)] \ \models\ \forall x\,[R(x) \Rightarrow \neg P(x)]\,.$$

Ce dernier syllogisme est bien une instance de CELARENT.

De même, le "C" de BOCARDO rappelle que ce dernier est réductible à BARBARA, par contradiction avec la majeure, donc en niant et permutant la majeure et la conclusion. L'inférence

$$\exists x\,[Q(x) \wedge \neg R(x)]\,,\ \forall x\,[Q(x) \Rightarrow P(x)] \ \models\ \exists x\,[P(x) \wedge \neg R(x)]$$

devient

$$\neg\exists x\,[P(x) \wedge \neg R(x)]\,,\ \forall x\,[Q(x) \Rightarrow P(x)] \ \models\ \neg\exists x\,[Q(x) \wedge \neg R(x)]\,,$$

c'est-à-dire

$$\forall x\,[P(x) \Rightarrow R(x)]\,,\ \forall x\,[Q(x) \Rightarrow P(x)] \ \models\ \forall x\,[Q(x) \Rightarrow R(x)]\,,$$

dans laquelle nous reconnaissons BARBARA. Le lecteur vérifiera sans peine, grâce au code "SMC", que les autres syllogismes valides se réduisent de manière analogue à la première figure.

6.4.7 Théorie ancienne de la réduction, syllogismes quasi-valides

Les scolastiques négligeaient en général les syllogismes "subalternes" BARBARI, CELARO, CAMESTROS, CESARO et CAMENOS ; le code "SMC" ne s'applique pas à eux, puisqu'on les rattache directement aux syllogismes valides dont ils proviennent, et non aux syllogismes valides de la première figure. En revanche, les "superalternes" DARAPTI, FELAPTON, BRAMANTIP et FESAPO étaient largement acceptés comme valides, et donc susceptibles d'être réduits au syllogisme valide de la première figure dont ils possèdent l'initiale.[87] Cette réduction mettait en jeu une quatrième méthode, codée par la lettre P, à savoir la *conversion partielle*,[88]

[87] Cette différence de point de vue entre la scolastique et la logique moderne n'indique pas la supériorité de la seconde sur la première ; la situation est plus nuancée et nous y reviendrons (brièvement) plus loin.

[88] ou "par accident", ou "limitative".

La conversion partielle n'est pas une opération valide au sens moderne. Elle consiste en la transformation de l'universelle affirmative

$$\forall x\,[A(x) \Rightarrow B(x)]$$

en la particulière affirmative

$$\exists x\,[B(x) \wedge A(x)]\,.$$

Il s'agit donc d'une subalternation (qui n'est que "quasi-valide"), suivie d'une conversion simple (qui est valide). On voit que si on applique la conversion partielle à la mineure de DARAPTI et de FELAPTON, ceux-ci deviennent respectivement DARII et FERIO. La même opération, accompagnée d'une conversion simple de la majeure, transforme FESAPO en FERIO. Enfin, BRAMANTIP devient BARBARA par permutation des prémisses et conversion partielle inverse de la conclusion.[89]

6.4.8 Critique de la méthode classique de réduction

Une première critique possible de la notion classique de réduction est qu'elle favorise la non-distinction entre validité et quasi-validité. Cette distinction est naturellement essentielle dans le cadre de la logique mathématique (où seule la validité a droit de cité) mais la théorie du syllogisme n'est pas une théorie mathématique, et lui appliquer des critères mathématiques de qualité ne serait pas pertinent.[90] En fait, dans la mesure où le langage naturel était jusqu'il y a peu le seul véhicule du raisonnement, comment s'étonner de l'acceptation d'une prémisse implicite puisque, dans le discours quotidien, cette prémisse est (quasi) toujours vérifiée ? On peut arguer que dans ce contexte, le problème de la non-validité de la subalternation ne méritait qu'une note en bas de page.

La notion même de subalternation suscite une deuxième critique : pourquoi recourir à la conversion par accident, opération plus complexe, plutôt qu'à la subalternation, pour la réduction des syllogismes quasi-valides des troisième et quatrième figures ? La réponse est analogue à la précédente. La théorie classique est liée au langage naturel, et dans ce langage la conversion par accident est plus naturelle, c'est-à-dire plus usitée, que la subalternation. Conclure, du fait que (tous) les Espagnols sont des Européens, que certains Espagnols sont des Européens est le plus souvent sans intérêt. Au contraire, l'administration européenne déduira avec profit, du même fait, que certains Européens sont des Espagnols, puisque cela entraîne la nécessité de traduire en espagnol les documents européens.

On peut enfin s'inquiéter de ce que, sans grande justification, c'est en sens inverse que la conversion par accident est utilisée lors de la réduction de BRAMANTIP à BARBARA. Cette troisième objection est difficile à réfuter formellement dans le cadre d'une théorie qui, justement, n'est pas formelle, mais on peut néanmoins observer que, lors de l'application d'une opération non symétrique (c'est-à-dire différente de l'opération inverse), il convient de distinguer les prémisses de la conclusion. Nous avons déjà observé que la subalternation de la conclusion, mais aussi la superalternation d'une prémisse, transforme un syllogisme

[89] Autrement dit, si on permute les prémisses de BARBARA et que l'on opère la conversion partielle sur la conclusion, on obtient BRAMANTIP.

[90] En revanche, revoir cette théorie, et en particulier la réduction, sous l'angle mathématique est pertinent, comme nous le verrons un peu plus loin.

valide en syllogisme quasi-valide ; il n'est donc pas surprenant de noter que la conversion par accident de la conclusion de BARBARA (accompagnée d'une transposition des prémisses) conduit à BRAMANTIP, alors que c'est l'opération inverse, appliquée cette fois à une prémisse (la mineure) qui transforme DARII en DARAPTI.

C'est une objection moins grave, mais plus fondée que les précédentes, qui pourra nous inciter à revoir, du moins dans sa présentation, le problème de la réduction des syllogismes. La solution classique recourt à quatre méthodes de réduction, dont nous ne remettons plus en doute le bien-fondé. Il est cependant paradoxal que l'application de ces méthodes ne soit pas librement autorisé. Par exemple, nous savons que l'on passe de BARBARA à BARBARI par subalternation de la conclusion mais, si l'on choisit plutôt de convertir partiellement cette conclusion (opération dont nous venons de souligner le caractère naturel), nous obtenons un syllogisme, dont le nom traditionnel est BARALIPTON, qui n'apparaît pas dans notre tableau récapitulatif. Les inconvénients concrets de cette lacune, et de toute autre du même ordre, sont très limités ; en particulier, BARALIPTON peut être vu comme "synonyme" ou comme "variante", de BARBARI, puisqu'une conversion simple de la conclusion permet de passer de l'un à l'autre. Sans doute, mais BARALIPTON peut tout aussi bien être considéré comme une variante de BRAMANTIP, puisqu'une transposition des prémisses permet de passer de l'un à l'autre ! La question n'est pas de savoir quelle figure, la première ou la quatrième, peut revendiquer notre intrus ; en fait, celui-ci montre clairement que la répartition des syllogismes en figures est arbitraire.

Remarque historique. Au départ, la théorie du syllogisme ne prévoyait pas la quatrième figure. Certains logiciens du Moyen-Age, peut-être par respect pour le travail d'Aristote,[91] ont refusé d'introduire celle-ci, préférant une "première figure bis", composée des "modes indirects" de la première figure. Le mode indirect est celui où la conclusion est de type *PR* et non de type *RP*. Cela est donc en contradiction avec la règle traditionnelle selon laquelle, dans la conclusion d'un syllogisme, le mineur est "sujet" (vient en premier lieu) et le majeur est "prédicat" (vient en second lieu). Voici le tableau général des syllogismes, obtenu en ajoutant cette "nouvelle" figure :

Figure :	première	première bis	deuxième	troisième	quatrième
Majeure	QR	QR	RQ	QR	RQ
Mineure	PQ	PQ	PQ	QP	QP
Conclusion	PR	RP	PR	PR	PR

On observe immédiatement que la figure supplémentaire est bien un substitut acceptable de la quatrième figure, comme le souhaitaient ses promoteurs. Plus précisément, un syllogisme de la première figure bis est valide si et seulement si en permutant ses prémisses (ce qui est sans effet logique) on obtient un syllogisme valide de la quatrième figure. La même règle est valable pour les syllogismes quasi-valides. Pour s'en convaincre, il suffit d'effectuer l'opération de permutation ; elle transforme le syllogisme de première figure bis

$$QR, PQ \models RP$$

en

$$PQ, QR \models RP$$

dans lequel on reconnaît le type de la quatrième figure, dans lequel les rôles de P et de R ont été échangés. Dans la mesure où les modes valides et quasi-valides de la quatrième figure sont

AEE-4, IAI-4, EIO-4 (valides) ; aeo-4, aai-4, eao-4 (quasi-valides) ,

[91] Respect mal placé, qui consiste à figer le défaut d'un travail par ailleurs remarquable au lieu de le corriger . . .

les modes valides et quasi-valides de la première figure bis seront

EAE-1', AII-1', IEO-1' (valides) ; eao-1', aai-1', aeo-1' (quasi-valides) .

Leurs noms traditionnels sont CELANTES, DABITIS, FRISESOMORUM et CEMANOS ?, BARALIPTON, FAPESMO.[92]

Une classification présentant un certain degré d'arbitraire n'est pas nécessairement une mauvaise classification mais, plus qu'une autre, elle est susceptible d'être remise en question et/ou de donner lieu à des variantes qui, au cours du temps, conduisent à une certaine confusion.[93]

Si nous considérons fondée cette troisième objection, il conviendrait en principe d'améliorer le système. Cela peut se faire de plusieurs manières, plus ou moins minimalistes ou extensives ; nous esquissons d'abord ici une manière minimaliste. Tout d'abord, nous pouvons conserver les notions de figures et de modes, mais nous devons admettre que leur rôle est plus taxinomique et mnémotechnique que réellement "logique". Pour la notion de mode indirect, nous pouvons la rejeter, ou au contraire l'admettre pour les quatre figures. La première solution aurait l'avantage de nous limiter à 256 syllogismes possibles, dont 15 valides et 9 quasi-valides ; elle nous dispenserait aussi de chercher des noms supplémentaires ! La seconde solution aurait l'avantage, auquel l'algébriste qui sommeille en tout logicien contemporain serait sensible, de rendre l'ensemble des 512 syllogismes, celui des 30 syllogismes valides et celui des 18 syllogismes quasi-valides fermés pour les opérations S, M et C.[94] De même, l'opération P, ainsi que l'opération inverse, la subalternation et la superalternation font passer d'un syllogisme valide à un syllogisme quasi-valide et réciproquement.

6.4.9 Une méthode moderne de réduction

Les deux modifications minimalistes que nous venons d'envisager sont inspirées par l'algèbre, discipline essentielle en logique formelle contemporaine.[95] L'algébriste aime que ses opérations soient internes (les opérandes et le résultat de l'opération appartiennent à un même ensemble, fixé à l'avance), mais aussi partout définies (tous les éléments de l'ensemble sont des opérandes appropriés et fournissent toujours un résultat). L'algébriste est aussi friand de formes canoniques, auxquelles peuvent se ramener tous les éléments de l'ensemble ; cela signifie que le problème de la réduction reste un problème très moderne. Réduire par exemple BOCARDO à BARBARA est a priori aussi justifié que réduire le polynôme $(x-3)(x^2+1)+3x(x-1)+5-x$ à la forme habituelle x^3-3x+2, ou à la forme factorisée $(x-1)^2(x+2)$. Une bonne forme canonique doit être restrictive, de telle sorte que l'ensemble des objets canoniques soit aussi réduit et aussi simple que possible, mais elle doit rester suffisamment générale pour que tout objet puisse s'y ramener.

[92]Comme dans la quatrième figure, la tradition semble avoir omis l'un des syllogismes quasi-valides.

[93]Les noms traditionnels des syllogismes ne sont pas non plus uniques ; certains admettent des variantes. Par exemple, les noms BARBARI et BARALIPTON sont parfois confondus ou échangés ; BAMALIP est parfois employé pour BRAMANTIP ou pour BARALIPTON. Il est instructif de noter que la confusion croît avec l'arbitraire ; les syllogismes valides des trois premières figures n'ont qu'une appellation. Les syllogismes quasi-valides et ceux de la quatrième figure en reçoivent plusieurs parce que l'on hésite à les mentionner. Dans le même ordre d'idée, le code SMCP fonctionne pour les quasi-valides superalternes, mais pas pour les quasi-valides subalternes, d'ailleurs absents de la plupart des listes. Ce défaut peut (doit ?) suggérer que l'on modifie leurs noms.

[94]Cela signifie que si l'on applique l'une de ces opérations à un syllogisme de l'un de ces ensembles, on obtient un syllogisme du même ensemble.

[95]La logique mathématique est souvent considérée comme une branche de l'algèbre.

Dans le cas du syllogisme, la situation n'est pas mauvaise, puisqu'il n'y a que quatre syllogismes canoniques. Cependant, on peut envisager de réduire ce nombre. D'autre part, l'opération de conversion, par exemple, ne s'applique pas à tout composant (prémisse ou conclusion) d'un syllogisme, puisque seules les affirmatives particulières et les négatives universelles peuvent faire l'objet d'une conversion simple. Enfin, l'ensemble des 512 syllogismes possibles pourrait lui-même être agrandi.[96]

Tenter de réduire à BARBARA les trois autres syllogismes canoniques CELARENT, DARII et FERIO nous permettra de résoudre toutes ces imperfections simultanément. Le syllogisme CELARENT montre la nécessité d'inclure dans notre arsenal de réduction l'obversion, opération introduite au paragraphe 3.5.1, qui d'ailleurs était déjà bien connue de la scolastique. Appliquée au majeur, l'obversion transforme

$$\forall x\,[Q(x) \Rightarrow \neg R(x)]\,,\ \forall x\,[P(x) \Rightarrow Q(x)] \ \models\ \forall x\,[P(x) \Rightarrow \neg R(x)]$$

en

$$\forall x\,[Q(x) \Rightarrow \overline{R}(x)]\,,\ \forall x\,[P(x) \Rightarrow Q(x)] \ \models\ \forall x\,[P(x) \Rightarrow \overline{R}(x)]\,.$$

Pour réduire DARII à BARBARA, on peut d'abord appliquer la méthode de contradiction, avec la prémisse mineure ; l'inférence

$$\forall x\,[Q(x) \Rightarrow R(x)]\,,\ \exists x\,[P(x) \wedge Q(x)] \ \models\ \exists x\,[P(x) \wedge R(x)]$$

devient

$$\forall x\,[Q(x) \Rightarrow R(x)]\,,\ \forall x\,[P(x) \Rightarrow \neg R(x)] \ \models\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\,.$$

Par obversion sur le (nouveau) moyen R, on obtient

$$\forall x\,[Q(x) \Rightarrow \neg\overline{R}(x)]\,,\ \forall x\,[P(x) \Rightarrow \overline{R}(x)] \ \models\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\,;$$

la première prémisse peut maintenant faire l'objet d'une conversion (simple), ce qui donne

$$\forall x\,[\overline{R}(x) \Rightarrow \neg Q(x)]\,,\ \forall x\,[P(x) \Rightarrow \overline{R}(x)] \ \models\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\,;$$

enfin une seconde obversion, sur Q, donne

$$\forall x\,[\overline{R}(x) \Rightarrow \overline{Q}(x)]\,,\ \forall x\,[P(x) \Rightarrow \overline{R}(x)] \ \models\ \forall x\,[P(x) \Rightarrow \overline{Q}(x)]\,,$$

inférence dans laquelle nous reconnaissons BARBARA. Remarquons par ailleurs qu'une simple obversion permet de réduire FERIO à DARII ; cela suffit à montrer que tout syllogisme valide est réductible à BARBARA, éventuellement via CELARENT, DARII ou FERIO.

[96]Agrandir son domaine est toujours agréable au scientifique, et en particulier à l'algébriste, qui au cours des siècles est passé des entiers naturels aux entiers relatifs, puis aux nombres rationnels, et enfin aux nombres réels et aux nombres complexes ; en cours de route, on a gagné des propriétés intéressantes (densité, complétude, clôture algébrique) sans perdre beaucoup de propriétés essentielles. (Notons quand même que l'ordre numérique n'est bien fondé que dans ℕ et qu'il n'existe plus dans ℂ.)

De manière analogue, tout syllogisme quasi-valide peut se ramener à BARBARI, en faisant usage des mêmes transformations.[97] Une simple obversion réduit CELARO à BARBARI. Pour ramener BRAMANTIP à BARBARI, il suffit de permuter les deux prémisses et d'opérer une conversion simple sur la conclusion. La réduction de FESAPO nécessite l'application de la méthode de contradiction, portant sur la majeure ; on passe ainsi de

$$\forall x\,[R(x) \Rightarrow \neg Q(x)]\,,\ \forall x\,[Q(x) \Rightarrow P(x)]\ \models\ \exists x\,[P(x) \wedge \neg R(x)]$$

à

$$\forall x\,[P(x) \Rightarrow R(x)]\,,\ \forall x\,[Q(x) \Rightarrow P(x)]\ \models\ \exists x\,[R(x) \wedge Q(x)]\,.$$

Le cas de FELAPTON, qui se ramène à FESAPO par conversion simple de la majeure, est réglé de même, et donc aussi celui de DARAPTI, qu'une obversion sur le majeur réduit à FELAPTON.

Par ailleurs, CESARO se réduit à CELARO par conversion simple de la majeure ; CAMESTROS se réduit à CESARO par obversion du moyen, et enfin CAMENOS se réduit à CAMESTROS par conversion simple de la mineure ; tous ces syllogismes quasi-valides se réduisent donc aussi à BARBARI.

Cela s'applique naturellement aussi au mode indirect ; à titre d'exemple, considérons FAPESMO :

$$\forall x\,[Q(x) \Rightarrow R(x)]\,,\ \forall x\,[P(x) \Rightarrow \neg Q(x)]\ \models\ \exists x\,[R(x) \wedge \neg P(x)]\,.$$

Par contradiction sur la mineure, on a

$$\forall x\,[Q(x) \Rightarrow R(x)]\,,\ \forall x\,[R(x) \Rightarrow P(x)]\ \models\ \exists x\,[P(x) \wedge Q(x)]$$

qui par mutation devient

$$\forall x\,[R(x) \Rightarrow P(x)]\,,\ \forall x\,[Q(x) \Rightarrow R(x)]\ \models\ \exists x\,[P(x) \wedge Q(x)]\,;$$

enfin, par conversion simple de la conclusion, on obtient

$$\forall x\,[R(x) \Rightarrow P(x)]\,,\ \forall x\,[Q(x) \Rightarrow R(x)]\ \models\ \exists x\,[Q(x) \wedge P(x)]\,,$$

c'est-à-dire BARBARI.

A ce stade, on peut "oublier" figures et modes et dire qu'un syllogisme est valide si et seulement si on peut le ramener à BARBARA par conversion simple, contradiction, mutation (permutation des prémisses) et obversion. De même, un syllogisme est quasi-valide si et seulement si les mêmes opérations permettent de le ramener à BARBARI. Enfin, un syllogisme est valide si et seulement si la subalternation de sa conclusion universelle ou la superalternation d'une prémisse existentielle le rend quasi-valide et un syllogisme est quasi-valide si et seulement si la subalternation d'une prémisse universelle ou la superalternation de sa conclusion existentielle le rend valide.

Cette simplification drastique de la théorie ne résout pas tout. D'abord, la conversion simple ne s'applique toujours qu'aux universelles négatives et aux particulières affirmatives ; en outre, l'obversion ne s'applique qu'au second terme (le prédicat) des prémisses et des conclusions. Le remède "algébrique" est simple : il suffit d'étendre les notions de syllogisme et de conversion simple de manière à ce que cette opération, de même que l'obversion, s'appliquent sans restriction. Pour cela, il faut

[97] La conversion partielle reste naturellement interdite.

– admettre que le terme sujet puisse être nié ;
– généraliser la conversion simple sur les particulières en la règle de commutativité de la conjonction ;
– généraliser la conversion simple sur les particulières en la règle de contraposition de l'implication.

Le tableau ci-dessous résume la "conversion généralisée" :

$$\begin{aligned}
\forall x\,(P(x) \Rightarrow Q(x)) &\simeq \forall x\,(\neg Q(x) \Rightarrow \neg P(x))\,;\\
\forall x\,(P(x) \Rightarrow \neg Q(x)) &\simeq \forall x\,(Q(x) \Rightarrow \neg P(x))\,;\\
\forall x\,(\neg P(x) \Rightarrow Q(x)) &\simeq \forall x\,(\neg Q(x) \Rightarrow P(x))\,;\\
\forall x\,(\neg P(x) \Rightarrow \neg Q(x)) &\simeq \forall x\,(Q(x) \Rightarrow P(x))\,;\\
\exists x\,(P(x) \wedge Q(x)) &\simeq \exists x\,(Q(x) \wedge P(x))\,;\\
\exists x\,(P(x) \wedge \neg Q(x)) &\simeq \exists x\,(\neg Q(x) \wedge P(x))\,;\\
\exists x\,(\neg P(x) \wedge Q(x)) &\simeq \exists x\,(Q(x) \wedge \neg P(x))\,;\\
\exists x\,(\neg P(x) \wedge \neg Q(x)) &\simeq \exists x\,(\neg Q(x) \wedge \neg P(x))\,.
\end{aligned}$$

Naturellement, admettre que le terme-sujet puisse être nié fait passer de 8 à 16 le nombre de $\{P, Q\}$-formules, et donc de $8^3 = 512$ à $16^3 = 4\,096$ le nombre de syllogismes ; cela n'est en rien gênant, même si le gain n'est guère utile en pratique.

6.4.10 Le point de vue moderne

La méthode des tableaux sémantiques et d'autres méthodes générales permettent, sans grand effort, d'identifier les syllogismes valides et ceux qui ne le sont pas. En fait, des techniques plus expéditives existent pour le calcul des prédicats monadiques, dont la théorie du syllogisme n'est qu'un fragment. La théorie du syllogisme est à la théorie générale ce que la théorie de Kepler est à celle de Newton : un cas particulier, mais très important (cf. § 1.1.4). Il ne faudrait pas en conclure que la théorie classique n'a plus d'intérêt. En effet, la brève présentation que nous venons de tracer est très partielle. Elle omet notamment les difficultés liées à l'emploi des langues naturelles, qui permettent d'écrire des phrases dont l'interprétation (et a fortiori la modélisation) n'est pas évidente. Il faut aussi remarquer que, si on les enchaîne, les syllogismes permettent d'élaborer des raisonnements compliqués. Une importante proportion des raisonnements courants sont de ce type.

D'un point de vue didactique, la théorie classique est aussi pour nous une occasion de réfléchir, en utilisant l'éclairage de la théorie moderne. Nous n'avons fait qu'effleurer ce point en traitant de manière moderne la question de la réduction, et en généralisant quelque peu la notion de syllogisme. La théorie moderne permet aussi d'éclairer certains points confus ou controversés, comme l'opportunité d'introduire une quatrième figure ou de préférer les modes indirects de la première figure ; dans le même ordre d'idée, on a pu préciser la distinction entre syllogismes valides et quasi-valides et montrer que la partition de ces derniers en subalternes (souvent omis dans la nomenclature) et superalternes (habituellement repris dans la nomenclature sous les noms DARAPTI, FELAPTON, BRAMANTIP et FESAPO) n'était pas logiquement justifiée.

Remarque. La prudence reste de mise sur ce dernier point. Apporter un éclairage nouveau à une controverse ne signifie pas la trancher. En particulier, on peut donner une justification extra-logique, mais néanmoins pertinente, à la partition traditionnelle. La scolastique notait déjà que, dans une situation

où les prémisses de BARBARA sont supposées vraies, on préférerait déduire la conclusion superalterne, plus riche, c'est-à-dire plus informative, que la conclusion subalterne. En revanche, dans une situation où les prémisses de DARAPTI, FELAPTON ou FESAPO sont supposées vraies, on n'a pas la possibilité de déduire autre chose que les conclusions correspondantes ; en particulier, déduire les superalternes de celles-ci serait incorrect. Le cas de BRAMANTIP est intermédiaire ; si l'on suppose vraies ses prémisses, il est possible de déduire "plus" que sa conclusion, mais à condition de changer de figure ; on passe de la quatrième à la première. A l'époque actuelle où les mécanismes de raisonnement sont vus comme des "producteurs d'information" et où un rendement maximal est systématiquement recherché, cette justification est intéressante.

Il est naturel qu'une théorie générale, comme la logique prédicative ou la théorie de Newton, supplante une théorie particulière, comme la syllogistique ou la théorie de Kepler. Néanmoins, quand un problème peut être résolu dans le cadre d'une théorie particulière, la résolution est souvent plus simple que dans un cadre général. Ce point justifie le maintien de théories particulières.

Dans notre contexte, la méthode des diagrammes de Venn, facile à comprendre et à mettre en œuvre, a en contrepartie l'inconvénient d'une portée très limitée par rapport à la méthode des tableaux sémantiques, par exemple. Il n'empêche que les diagrammes de Venn ne sont pas limités aux syllogismes. Pour le voir, considérons par exemple le raisonnement

$$\frac{\exists x\, P(x)\,,\ \forall x\,(P(x) \Rightarrow [Q(x) \Rightarrow R(x)])\,,\ \forall x\,[S(x) \vee Q(x)]\,,\ \forall x\,[R(x) \Rightarrow S(x)]}{\exists x\, S(x)}$$

Ce raisonnement est intuitivement correct. Si $P(a)$ est vrai, alors $Q(a)$ est faux ou $R(a)$ est vrai. Dans les deux cas, $S(a)$ est vrai. La méthode des diagrammes de Venn peut s'appliquer, mais quatre ensembles de base seront nécessaires au lieu de trois, un pour chacun des prédicats P, Q, R et S. Il y aura donc seize zones élémentaires, pour lesquelles nous adoptons une notation simple, qu'un exemple suffit à préciser. La zone $p\overline{q}rs$ contient les objets qui rendent P, R et S vrais et Q faux. La réunion des seize zones se note V. La zone contenant les objets qui rendent P vrai et Q faux (et qui attribuent une valeur de vérité quelconque à R et S) se notera $p\overline{q}$; on a donc

$$p\overline{q} = p\overline{q}rs \cup p\overline{q}r\overline{s} \cup p\overline{q}\overline{r}s \cup p\overline{q}\overline{r}\overline{s}\,.$$

En utilisant la notation ensembliste, notre raisonnement peut se récrire

$$\frac{p \neq \emptyset\,,\ pq\overline{r} = \emptyset\,,\ s \cup q = V\,,\ r\overline{s} = \emptyset}{s \neq \emptyset}$$

Le calcul ensembliste élémentaire suffit à montrer la validité de ce raisonnement :

1.	$p \neq \emptyset$	5.	$p\overline{q} \cup pqr \neq \emptyset$	$[1,2]$
2.	$pq\overline{r} = \emptyset$	6.	$ps \cup pqr \neq \emptyset$	$[3,5]$
3.	$s \cup q = V$	7.	$ps \cup pqs \neq \emptyset$	$[4,6]$
4.	$r\overline{s} = \emptyset$	8.	$s \neq \emptyset$	$[7]$

La théorie du syllogisme est sans doute très limitée mais, dans les limites de son champ d'action, elle est "parfaite" en ce sens qu'il est possible, sur base de cette théorie, de décider "mécaniquement", par l'emploi aveugle de règles simples et non ambiguës, si un syllogisme est valide, quasi-valide ou non valide. Si nous étendons la portée de la théorie, nous devons néanmoins essayer de conserver cette précieuse caractéristique, que l'on appelle la *décidabilité*.

Le calcul ensembliste élémentaire utilisé ici est-il décidable ? La réponse est affirmative, mais nous ne faison qu'esquisser une démonstration.[98] Supposons que n ensembles (donc n prédicats, tous monadiques) soient en jeu. Il existe 2^n zones élémentaires, et 2^{2^n} configurations, pour lesquelles chaque zone élémentaire est vide ou non. Il suffit de passer en revue toutes les configurations rendant vraies les prémisses et de vérifier qu'elles rendent vraie aussi la conclusion.[99]

Il faut aussi montrer que toute la logique des prédicats monadiques est réductible au calcul ensembliste élémentaire. Ce point n'est pas évident a priori ; considérons par exemple les deux formules

$$\exists x\, \forall y\, (\forall z\, [(P(z) \vee Q(y)) \Rightarrow R(a)] \Rightarrow \exists z\, [P(x) \wedge S(y) \wedge Q(z)])\,;$$

$$\forall x\, \exists y\, (\forall z\, [(P(z) \vee Q(y)) \Rightarrow R(a)] \Rightarrow \exists z\, [P(x) \wedge S(y) \wedge Q(z)])\,.$$

La difficulté liée aux variables distinctes et aux quantifications imbriquées n'est qu'apparente, en vertu d'équivalences logiques telles que

$$\begin{aligned} \forall x\, \exists y\, [A(x) \wedge B(y)] \ &\simeq\ [\forall x\, A(x) \wedge \exists y\, B(y)]\,, \\ \exists x\, \forall y\, [A(y) \Rightarrow B(x)] \ &\simeq\ [\exists y\, A(y) \Rightarrow \exists x\, B(x)]\,. \end{aligned}$$

La présence de constantes ne pose pas de problème non plus car celles-ci s'intègrent dans le calcul ensembliste élémentaire. (Elles peuvent aussi s'éliminer dans le cadre des formules elles-mêmes.) Les formules données plus haut se réduisent respectivement à

$$[(\exists x\, P(x)\ \vee\ \forall x\, Q(x)) \Rightarrow R(a)]\ \Rightarrow\ [\exists x\, P(x)\ \wedge\ \forall x\, S(x)\ \wedge\ \exists x\, Q(x)]\,;$$

$$[(\exists x\, P(x)\ \vee\ \exists x\, Q(x)) \Rightarrow R(a)]\ \Rightarrow\ [\forall x\, P(x)\ \wedge\ \exists x\, S(x)\ \wedge\ \exists x\, Q(x)]\,.$$

[98] Une vraie démonstration exigerait d'abord que nous décrivions formellement ce calcul, ce qui allongerait un peu trop cette digression . . .

[99] Cette procédure de décision est l'analogue, pour la logique des prédicats monadiques, de la méthode des tables de vérité. Comme cette dernière, elle est extrêmement inefficace mais de meilleures méthodes existent, dont l'informatique permet la mise en œuvre concrète.

Références

[1] I.M. Copi, *Introduction to logic*, Collier-MacMillan Publ., London, 1972 (quatrième édition).

[2] H.C.M. De Swart, *Logic : Mathematics, Language and Philosophy*, Peter Lang Verlag, Frankfurt, 1993.

[3] P. Gochet et P. Gribomont, *Logique, méthodes pour l'informatique fondamentale*, Hermès, Paris, 1989, 1998 (troisième édition).

[4] P. Gochet et P. Gribomont, *Logique, méthodes pour l'étude des programmes*, Hermès, Paris, 1994.

[5] P. Gochet, P. Gribomont et A. Thayse, *Logique, méthodes pour l'intelligence artificielle*, Hermès, Paris, 2000.